

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 19:04:01 ; Search time 3885 Seconds

(without alignments)
2344,808 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962
Sequence: 1 MATDVNSKNIAVQAQKIL.....KPHLOKLCGAINMLDEBN 188

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 100 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=x1h
-Q/cgna2.1/USPTO.spool.b/US10627571/runat.27072005.154720.27503/app.query.fasta_1.127
-DB=GenEmbl -QFMT=fasta -SUFFIX=rge -MINMATCH=0 -LOOPEXT=0 -LOOPEXT=0
-INITs-bits -START=1 -END=1 -MATRIX=blowm62 -TRANS=human40.cdt -LIST=100
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10627571.@CGN 1.1.3731.@runat.27072005.154720.27503 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DISPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_hcg:.*
3: gb_in:.*
4: gb_ov:.*
5: gb_ov:.*
6: gb_ph:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vl:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	954	99.2	1892	9 AF070671
2	954	99.2	1921	6 BD156880
3	954	99.2	1921	6 AX877846
4	954	99.2	1921	9 AK001931

5	950	98.8	573	9 CR457137	CR457137 Homo sapi
6	950	98.8	1729	6 BD156785	BD156785 primer fo
7	950	98.8	1729	6 AX877673	AX877673 Sequence
8	950	98.8	1729	9 AK001850	AK001850 Homo sapi
9	950	98.8	1814	9 BC005352	BC005352 Homo sapi
10	950	98.8	1915	9 AF099935	AF099935 Homo sapi
11	950	98.8	1988	6 CQ726075	CQ726075 Sequence
12	950	98.8	2003	9 AF099936	AF099936 Homo sapi
13	950	98.8	2081	9 BC007014	BC007014 Homo sapi
14	950	98.8	112626	9 AC035144	AC035144 Homo sapi
15	950	98.8	156277	9 AC026795	AC026795 Homo sapi
16	950	98.8	158057	9 AC138612	AC138612 Homo sapi
17	948	98.5	1943	6 BD13356	BD13356 110 human
18	944	98.1	2502	9 AK097284	AK097284 Homo sapi
19	920	95.6	1728	10 BC009090	BC009090 Mus muscu
20	920	95.6	184327	10 AC120859	AC120859 Mus muscu
21	918	95.4	237561	2 AC095257	AC095257 Rattus no
22	916	95.2	816	6 BD149395	BD149395 primer fo
23	916	95.2	816	6 AX869333	AX869333 Sequence
24	880	91.5	2099	5 AJ720906	AJ720906 Gallus ga
25	856.5	89.0	1766	9 AK097884	AK097884 Homo sapi
26	844.5	87.8	1559	5 CR760636	CR760636 Xenopus t
27	842.5	87.6	1193	5 BC072904	BC072904 Xenopus l
28	698	72.6	3986	5 BC052765	BC052765 Danio rer
29	668	72.6	205949	5 BX649252	BX649252 Zebrafish
30	666	69.2	1589	5 AJ720336	AJ720336 Gallus ga
31	651	67.7	1248	6 CQ736656	CQ736656 Sequence
32	651	67.7	1986	6 CQ841622	CQ841622 Sequence
33	651	67.7	1986	9 AK123281	AK123281 Homo sapi
34	651	67.7	188937	2 AC012678	AC012678 Homo sapi
35	651	67.7	189796	9 AC073964	AC073964 Homo sapi
36	633	65.8	110000	2 AC115187_1	Continuation (2 of
37	633	65.8	256285	2 AC115505	AC115505 Rattus no
38	633	65.8	273225	2 AC129440	AC129440 Rattus no
39	626	65.1	164258	2 AC118938	AC118938 Mus muscu
40	626	65.1	195574	2 AC107848	AC107848 Mus muscu
41	607	62.4	1938	5 BC076797	BC076797 Xenopus l
42	600	62.4	134497	5 BC053167	BC053167 Danio rer
43	600	62.4	189797	5 BX927313	BX927313 Zebrafish
44	593	61.6	32369	9 AC005339	AC005339 Homo sapi
45	593	61.4	22209	9 BC017672	BC017672 Homo sapi
46	590	61.3	561	6 CQ732643	CQ732643 Sequence
47	587	61.0	587	6 BD149633	BD149633 primer fo
48	587	61.0	587	6 AX869571	AX869571 Sequence
49	574	59.7	1357	5 BC053238	BC053238 Danio rer
50	574	59.4	107568	2 AC144398	AC144398 Rattus no
51	571	59.4	230310	2 AC120079	AC120079 Rattus no
52	567	58.9	2307	10 BC032199	BC032199 Mus muscu
53	567	58.9	180238	2 AC073688	AC073688 Mus muscu
54	567	58.9	222610	10 AC026385	AC026385 Mus muscu
55	567	58.9	232190	2 AC074167	AC074167 Mus muscu
56	556	57.8	2175	5 BC061657	BC061657 Xenopus l
57	556	55.7	515	9 AF098933	AF098933 Homo sapi
58	536	54.8	1200	9 BC063014	BC063014 Homo sapi
59	527	54.8	1268	6 AX078255	AX078255 Sequence
60	527	54.8	3116	9 HSM804686	AL833373 Homo sapi
61	527	54.8	154526	2 AL592111	AL592111 Homo sapi
62	527	54.8	159148	2 AL592424	AL592424 Human DNA
63	526	54.7	181220	2 AC149856	AC149856 Papio anu
64	526	54.7	182245	2 AC149842	AC149842 Papio anu
65	526	54.7	207952	2 AC151020	AC151020 Callithrix
66	524	54.5	1196	9 AK027120	AK027120 Homo sapi
67	520	54.1	555	9 CR457375	CR457375 Homo sapi
68	519	54.0	149954	2 AC151645	AC151645 Desmodus n
69	517	53.7	208375	2 AC147191	AC147191 Oryzomys
70	515	53.5	217685	2 AC149677	AC149677 Bos tauru
71	514	53.4	1433	10 BC079019	BC079019 Rattus no
72	514	53.4	174234	10 AC117098	AC117098 Rattus no
73	511	53.1	1198	10 BC055879	BC055879 Mus muscu
74	511	53.1	148675	10 AC131769	AC131769 Mus muscu
75	511	53.1	205702	10 AC140190	AC140190 Mus muscu
76	511	53.1	206230	10 AC084272	AC084272 Mus muscu
77	511	53.1	206230	10 AC084272	Mus muscu

78	503	52.3	1171	9	AF271774	Homo sapi
79	484.5	50.4	2156	3	AK112519	Giona int
80	401	41.7	829	6	CQ580257	Sequence
81	401	41.7	1821	3	AY095033	Drosophila
82	383.5	39.9	4406	6	CQ580256	Sequence
83	383.5	39.9	175118	3	AC010842	Drosophila
84	383.5	39.9	188272	3	AC005639	Drosophila
85	383.5	39.9	295225	3	AE003461	Drosophila
86	358	37.2	69208	2	AC020466	Drosophila
87	222	30.4	714	6	BD146713	Primer to
88	222	30.4	714	6	AX866651	Sequence
89	232	30.4	1602	6	BD160707	Primer to
90	232	30.4	1602	6	AX884081	Sequence
91	232	30.4	1602	6	AK024161	Homo sapi
92	250	26.0	340	6	AX898564	Sequence
93	250	26.0	340	6	BD034097	Sequence
94	166	17.3	252	6	AX898571	Sequence
95	166	17.3	252	6	BD034104	Sequence
96	112	11.6	301130	1	AE016763	Escherich
97	106	11.0	1329	3	AK116070	Giona int
98	105.5	11.0	303414	1	AE015938	Clostridi
99	104.5	10.9	3408	8	AF378568	Eremothec
100	104.5	10.9	110000	8	AE016816_3	Continuation (4 of

ALIGNMENTS

RESULT 1
AF070671 1892 bp mRNA linear PRI 21-JUN-1999
LOCUS Homo sapiens TNF-induced protein GSG-1 mRNA, complete cds.
DEFINITION AF070671
ACCESSION AF070671
VERSION AF070671.1 GI:3978237

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1892)
AUTHORS Horrevorts,A.J., Fontijn,R.D., van Zonneveld,A.J., de Vries,C.J.,
ten Cate,J.W., and Pannekoek,H.
Vascular endothelial genes that are responsive to tumor necrosis
factor-alpha in vitro are expressed in atherosclerotic lesions,
including inhibitor of apoptosis protein-1, stannin, and two novel
genes

JOURNAL Blood 93 (10), 3418-3431 (1999)
MEDLINE 99252096
PUBMED 10233894

REFERENCE 2 (bases 1 to 1892)
AUTHORS Horrevorts,A.J.G., Fontijn,R.D. and Pannekoek,H.
Direct Submission

JOURNAL Submitted (05-JUN-1998) Biochemistry, Academic Medical Center,
Weibergdreef 15, Amsterdam 1105 AZ, The Netherlands
Location/Qualifiers

FEATURES

CDS

1..1892
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="endothelial"
/tissue_type="umbilical vein"
98..664
/codon_start=1
/product="TNF-induced protein GSG-1"
/protein_id="AAC83229.1"
/db_xref="GI:3978238"
/translation="MATDVNSKNLAVQAKILGKGVSKSIATTLIDTSSVLEDEL
YVTRREYTONKKEAKRIIKLITKTVKLAIIYNNQFNQDGLAMKPKKVKYQLAMT
VVSFPOVDYTPDRNVTSLINCEGKREMLHOIQRHLPAKSHGRNVVDFHPSDEFLAA
LYNPFNFRPHLOKLCDDGINKMLDEENI"

Alignment Scores:
Pred. No.:

9.11e-76 Length: 1892

Score: 954.00 Matches: 187
Percent Similarity: 99.47%
Best Local Similarity: 99.47%
Query Match: 99.17%
DB: 9
Gaps: 0

US-10-627-571-2 (1-188) x AF070671 (1-1892)

QY	1	MetAlAThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu	20
DB	98	ATGGCCACAGATGCTTTTAAATTCAAAACCGCGCCGTTCAGCACAAAAGAGATCTTG	157
QY	21	GLYsMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerGluVal	40
DB	158	GGTAAATAGTGTCCAAATCCATCCGCCACCTTAAATAGACAAAGTAGTAGG	217
QY	41	LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnIysGlnIaGluys	60
DB	218	CTGGATGAGCTCTTACAGAGTGACCGAGGAGTACCCAAAACAAAGAGGACGAG	277
QY	61	IysIleIysAsnLeuIleTyThrValIleIysLeuAlaIleLeuTyThrAsnGln	80
DB	278	ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTTTATGAAATATCAG	337
QY	81	PheAsnGlnAspGluLeuAlaLeuMetGluysPheIysValIhIsgIleuAla	100
DB	338	TTTAAATCAAGATGAGCTGATTCATTCAGAAATTTAAGAAAGATTCATCAGCTGCT	397
QY	101	MetThrValIalSerPheIhIsgIlnValAspTyThrPheAspArgAsnValIeuserArg	120
DB	398	ATGACCGTGTGATGATTCATCAGGTCGATTTACCTTTGACCGGAATGTGTTATCCAG	457
QY	121	LeuLeuAsnGluCyArgGluMetLeuIhIsgIlnIleIsgIlnArgIhIseuthrAlaIys	140
DB	458	CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCACCGCCACTGCAAG	517
QY	141	SerIhIsgIArgValAsnAsnValPheAspIhIspSerAspCysGluPheLeuAlaIa	160
DB	518	TCACATGACGGGTTAATATGTCCTTTGATCATTTTCAGATTGGAATTTTGGCTGCG	577
QY	161	LeuTyAsnProPheGluAsnIhIspProIhIseGlnIysLeuCyAspGlyIleAsn	180
DB	578	TGTATTAATCCCTTTGGGAATTTAAACCCCACTTACAAAACATAATGATGATCAAC	637
QY	181	IysMetLeuAspGluGluAsnIle	188
DB	638	AAATGTTGATGAAGAAACATA	661

RESULT 2
BD156880 1921 bp DNA linear PAT 17-JAN-2003
LOCUS BD156880
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION BD156880.1 GI:27862638
VERSION BD156880.1
KEYWORDS JP 2002191363-A/11723.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1921)
Oca,T., Isogai,T., Nishikawa,T., Haysashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 11723 09-JUL-2002;
HELIX RESEARCH INSTITUTE
OS Homo sapiens (human)
PN JP 2002191363-A/11723
PD 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTO, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
SAITO,
JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
KEIICHI NAGAI, TETSUJI OTSUKI

PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10', C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
(140) . (703) .

FEATURES
Source 1. .1921
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 9.25e-76 Length: 1921
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x BD156880 (1-1921)

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlyleu 20
DB 140 ATGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTCAGGACAAAAGAGATCTTG 199
QY 21 GlyMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 200 GGTAAATGGTGTCCAAATCCATCGCCACCACTTAAATACAGACAAAGTAGTAGTG 259
QY 41 LeuAspGluLeuThrArgValThrArgGluThrGlnAsnIlysgluAlaGluVal 60
DB 260 CTGATGAGCTCTACAGAGTACCAAGGAGTACCAACCAAAACAAGAGAGGAGAG 319
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuThrArgAsnGln 80
DB 320 ATCATCAAGAACTCTCAAGACAGATCATCACTGCGCATTTTAAAGAAATATACAG 379
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 380 TTTAAATCAAGATGAGCTACAGATTCATGAGAAATTTAAAGAAAGTTCATCAGCTTCT 439
QY 101 MetThrValIysSerPheHisGlnValAspThrThrPheAspArgAsnValLeuSerArg 120
DB 440 ATGACCGTGGTCACTTCATCAGTGCATTAATACCTTTGACCGGAATGTTATCCAGG 459
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 500 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTTCACCTGCCAG 559
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
DB 560 TCACATGACGCGGTTAAATATGCTTTGATCATTTTTCAGATTTGTAATTTTGGCTGCC 619
QY 161 LeuThrAsnProPheGluAsnPhelysProHisLeuGlnIlyleuLysCysAspGlyIleAsn 180
DB 620 TTGTATTAATCTTTTGGGAATTTTAAACCCCACTTCAAAACCTATGTATGTATCAAC 679
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 680 AAAATGTTGATGACAGAGAACATTA 703

RESULT 3
AX877846 1921 bp DNA linear PAT 17-DEC-2003
LOCUS AX877846
DEFINITION Sequence 12751 from Patent EP1074617.
ACCESSION AX877846
VERSION AX877846.1 GI:40032582
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Oca,T., Isogai,T., Nishikawa,T., Hayaehi,K., Saito,K., Yamamoto,U.,
Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 12751 07-FEB-2001;
Research Association for Biotechnology (JRB)

FEATURES
Source 1. .1921
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

CDS

140. .706
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89903.1"
/db_xref="GI:40032583"
/translation="MATDVNSKNLAVQAKKTIKRWVSKSATTLIDPSSSEVDEL
YRTREYTONKKAETIKILIKTVIKLILIRNPNODELAMEKFKKHQIANT
VSEFQVDYTFPBNVLSRLNECREMLHQIIOHLLPAKSHGRVNNVFDHPSDFPLAA
LYNPFQNFPHLQKLCGINKMDEENI"

ORIGIN

Alignment Scores:
Pred. No.: 9.25e-76 Length: 1921
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x AX877846 (1-1921)

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlyleu 20
DB 140 ATGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTCAGGACAAAAGAGATCTTG 199
QY 21 GlyMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 200 GGTAAATGGTGTCCAAATCCATCGCCACCACTTAAATACAGACAAAGTAGTAGTG 259
QY 41 LeuAspGluLeuThrArgValThrArgGluThrGlnAsnIlysgluAlaGluVal 60
DB 260 CTGATGAGCTCTACAGAGTACCAAGGAGTACCAACCAAAACAAGAGAGGAGAG 319
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuThrArgAsnGln 80
DB 320 ATCATCAAGAACTCTCAAGACAGATCATCACTGCGCATTTTAAAGAAATATACAG 379
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 380 TTTAAATCAAGATGAGCTACAGATTCATGAGAAATTTAAAGAAAGTTCATCAGCTTCT 439
QY 101 MetThrValIysSerPheHisGlnValAspThrThrPheAspArgAsnValLeuSerArg 120
DB 440 ATGACCGTGGTCACTTCATCAGTGCATTAATACCTTTGACCGGAATGTTATCCAGG 459
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 500 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTTCACCTGCCAG 559
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
DB 560 TCACATGACGCGGTTAAATATGCTTTGATCATTTTTCAGATTTGTAATTTTGGCTGCC 619
QY 161 LeuThrAsnProPheGluAsnPhelysProHisLeuGlnIlyleuLysCysAspGlyIleAsn 180
DB 620 TTGTATTAATCTTTTGGGAATTTTAAACCCCACTTCAAAACCTATGTATGTATCAAC 679
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 680 AAAATGTTGATGACAGAGAACATTA 703

RESULT 4
AK001931
LOCUS Homo sapiens CDNA FL111069 fis, clone PLACE1004930, highly similar
DEFINITION to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK001931
VERSION AK001931.1 GI:7023502
KEYWORDS oligo cloning; full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Houchi, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A.,
Sasaki, N., Aotsuma, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Nomiya, H., Satoh, N., Takani, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumaga, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togauchi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702035
REFERENCE
AUTHORS 2
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Aotsuma, S., Yoshikawa, Y.,
Matsunawa, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Sasaki, N.
NEDO human cDNA sequencing project
Unpublished
TITLE 3 (bases 1 to 1921)
JOURNAL Isogai, T. and Otsuki, T.
REFERENCE Direct Submission
AUTHORS Isogai, T. and Otsuki, T.
TITLE Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing and clone selection;
Helix Research Institute (supported by Japan Key Technology Center
etc.) and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
source
1..1921
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004930"

ORIGIN
Alignment Scores:
Pred. No.: 9,25e-76 Length: 1921
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: Gaps: 0
US-10-627-571-2 (1-188) x AK001931 (1-1921)
QY 1 MetAAThPaSPVaLPheAsnSerLySAsnLeuAValGlnAlGlnLySLeu 20
|||||
140 ATGGCCACAGAGTCTTAAATCCAAAACCTGGCCGTCAGGCAAAAAGAGATCTTG 199
DB 21 GLYSMetValSerLySerLLeaLThrThrLeuLeaSPThSerSerGluVal 40
|||||
200 GGTAAATGATGTCCTCAATCCATCCGCCACCTTATAGACACAAAGTAGTAGGTC 259
QY 41 LeuAPGluLeuTyArgValThrArgGluTyThrGlnAsnLysGlnAgiLys 60
|||||
260 CTGATGAGCTCTACAGAGTACAGGAGGTACACCAAAACAAGAGGAGAGAG 319
QY 61 LysLLeuAsnLeuLLeuThrValLLeuLLeuAlaLLeuTyArgAsnGln 80
|||||
320 ATCATCAAGAACCTCATCAAGCACTCATCAAGCTGCGCATCTTATAGAAATATCAG 379
DB 81 PheAsnGlnAPGluLeuAlaLeuMetGluTyPheLysLysValHisGlnLeuAla 100
|||||
380 TTTAATCAGATGAGCTACATGATGAGAAATTTAAGAGAAAGTTATCATGCTTCT 439
QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
|||||
440 ATGACCGTGTGATTCATCAGGTGATTAATCTTTGACCGGATGTGTATTCAGG 499
DB 121 LeuLeuAsnGluCyArgGluMetLeuHisGlnLLeuGlnArgHisLeuThrAlaLys 140
|||||
500 CTGTAAAGAAAGATGAGAGAGATGTGCAACCAATCATTCACGCGCATCTGCAAG 559
QY 141 SerHisGluArgValAsnAsnValPheAspHisPheSerArgSPGluPheLeuAlaLys 160
|||||
560 TCAATGACGAGGCTTATATATGTTTATCTTTTCAGATTGTAATTTTGGCTGCC 619
DB 161 LeuTyArgSPProPheGluAsnPhelYSPProHisLeuGlnLysLeuCyAspGlyLLeuAsn 180
|||||
620 TTGTATATATCTTTTGGGAATTTAAACCCCACTTACAAAACCTATGTGATGTATCAGC 679
QY 181 LysMetLeuAspGluGluAsnLLeu 188
|||||
680 AAAATGTTGATGAAGAGAAACATA 703
DB
RESULT 5
LOCUS CR457137 573 bp mRNA linear PRI 03-JUN-2004
DEFINITION Homo sapiens full open reading frame cDNA clone RZD0834H127D for
gene TNFAIP8, tumor necrosis factor, alpha-induced protein 8;
complete cds, incl. stopcodon.
ACCESSION CR457137
VERSION CR457137.1 GI:48146390
KEYWORDS Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 573)
AUTHORS Ebert, L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.
Cloning of human full open reading frames in Gateway(TM) system
entry vector (pDONR201)
JOURNAL Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

2 (bases 1 to 573)
Ebert, L., Schick, M., Neubert, P., Schatten, R., Henze, S. and Korn, B.
Direct Submission
Submitted (03-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer
Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,
Germany
RZPD: RZPD0834H127, ORFNO 1910
www.rzpd.de/cgi-bin/products/c1.cgi?cloneID=RZPD0834H127D RZPDLIB:
Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
834
www.rzpd.de/cgi-bin/products/showLib.pl.cgi?response?libNo=834
Contact: Inge Arlt
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
contact RZPD (customer.service@rzpd.de) for further information.
This CDS clone is a part of a collection of human full length
expression clones generated by RZPD.
This CDS has been inserted into pDONR201 via a BP Clonase(TM)
reaction. Additional sequence has been added in front of the start
codon: att. AAAA GCA GGC (ATG).
The last base of the last coding triplet has been changed to T,
which might lead to an amino acid change at the C terminus of the
polypeptide.
The stop codon has been set to TAA followed by
TTAACCAGCTTCTT. att. Compared to the reference sequence NM_014350
we did not find any amino acid exchanges.
Clone distribution: <http://www.rzpd.de/products/orfclones/>.
Location/Qualifiers

FEATURES

source

1..573
/organism="Homo sapiens"
/mol_type="cDNA"
/db_xref="taxon:9606"
/clone="RZPD0834H127D"
/clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
/lab_host="DH10B"
/note="Vector: pDONR201, Site_1: attP1, Site_2: attP2"
1..573
/gene="TNFAP1P8"
/gene="TNFAP1P8"
1..573
/gene="TNFAP1P8"
/codon_start=1
/protein_id="CAG33418.1"
/db_xref="GI:48146391"
/translation="MAVAVDVNSKLVLAQVQKKIIGKRVSKSIATTLIDTTSSEVLD
ELRVRYTQNKKEAKKIINKIKTVIKALIRNNQFQDILAMEKFKKQVQLA
MTVVSFHOVDYFDENVLSRLINCEKREMLHQIQRHLTAKSHGRVANVVDHFDSCBFL
AALYNPFNGKPKHLQKLCQKINMLDEBNIT"

ORIGIN

Alignment Scores:

Pred. No.: 6,04e-76 Length: 573
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x CR457137 (1-573)

QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnIlyGlyIleLeu 20
Db 7 GTGGCCACAGATGCTTAATTCAAAACCTGGCCCTTACGACACAAAAGAACTCTG 66
QY 21 GllYlVseMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
Db 67 GGTAAATGCTGTCCAATCATCGCACCACTTAATACGACACAACTAGTAGGCTG 126

QY 41 LeuAspGluLeuThrArgGluThrArgGluThrGlnAsnIysGluAlaGluIys 60
Db 127 CTGGATGAGCTCTTACAGAGTACAGAGGAGTACACCAACAAAGAGAGGAGAG 186
QY 61 lYsIlelYsAsnLeuIlelYsThrValIlelYsLeuAlaIlelYsThrArgAsnAngin 80
Db 187 ATCATCAGAGACTCATCAAGACAGCATCAAGCTGCGCATTTTATAGGAATATACG 246
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluIysPheIysIysValHisGlnLeuAla 100
Db 247 TTTAAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGAAATTCATCAGCTGCT 306
QY 101 MetThrValIysSerPheHisGlnValAspIlyThrPheAspArgAsnValLeuSerArg 120
Db 307 ATACCGTGCTGCTGCTTCCATCCATCGAGGANTTATACCTTTGACCGGAATGTTATCCAGG 366
QY 121 LeuLeuAsnGluIysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIys 140
Db 367 CTGTTAAATGAATGACAGAGAGAGTCTGCACCAATCATTCAGCGCCATCTGCACAG 426
QY 141 SerHisGluArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
Db 427 TCACATGACGCGGCTTAATATGTTGATGATTTTCAGATTGTGAATTTTGGCTGCC 486
QY 161 LeuThrAspProPheGluIysPheIysProHisLeuGluIlyLeuIysAspGluIleAsn 180
Db 487 TGTATATATCTCTTTGGGAATTTTAAACCCACTTACAAAACTAATGATGATATCAC 546
QY 181 lYsMetLeuAspGluIysAsnIle 188
Db 547 AAAATGTTGATGATGAAGAAACAT 570

RESULT 6
BD156785 1729 bp DNA linear PAT 17-JAN-2003
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BD156785 1729 bp DNA linear PAT 17-JAN-2003
BD156785.1 GI:27862543
JP 2002191363-A/11628.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 1729)
Ota, T., Iwagaki, T., Nishikawa, T., Hayaishi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 11628 09-JUL-2002;
HELIK RESEARCH INSTITUTE

COMMENT

OS Homo sapiens (human)
PN JP 2002191363-A/11628
PF 09-JUL-2002
PF 28-JUL-2000 JP 2000280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU
PI SAITO,
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUJI OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/ PC
10', C12P21/02, C12Q1/68//C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
FT CDS (71)..(664).

FEATURES

source
1..1729
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores: 1.89e-75 Length: 1729
Pred. No.: 1.89e-75

Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x BD156785 (1-1729)

```

QY      1 MetAlaThrAspValaPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
      :::
DB      101 GGGCCACAGATGCTTTTAATTCAAAACCTGGCCGTTCAAGCACAAAAGAGATCTTG 160
QY      21 GlyLysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      |||
DB      161 GGTAAATGCTGTCCAAATCCATCCGACCACTTAAATAGACACAAAGTAGTAGG 220
QY      41 LeuAspGluLeuTyrArgValaThrArgIuTyrThrGlnAsnIysIleGluAlaGluLys 60
      |||
DB      221 CTGGATGAGCTTACAGAGTGAACGAGGAGTACACCAAAACAAAGAGGAGAGAG 280
QY      61 LysIleLysAsnLeuIleIysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
      |||
DB      281 ATCATCAAGAACTCATCAAGACAGATCATCAAGCTGCCATCTTAAAGAAATTAATCAG 340
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysIleLysValaHisGlnLeuAla 100
      |||
DB      341 TTTAAATCAAGATGAGCTTGAATGAGGAAATTTAAAGAAAGATTCATCAGCTTGCT 400
QY      101 MetThrValaIleSerPheHisGlnValaAspTyrThrPheAspArgAsnValaLeuSerArg 120
      |||
DB      401 ATGACCGTGCTGATTCCATCAGTGATGATTATACCTTTGACCGGAATGTTATCCAG 460
QY      121 LeuLeuAsnGlnCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||
DB      461 CTGTAAATGAATGACAGAGATGCTGCACCAATATATTCAGGCCACCTCACTGCCAAG 520
QY      141 SerHisGlyArgValaAsnAsnValaPheAspHisPheSerAspCysGluPheLeuAlaAla 160
      |||
DB      521 TCACATGACGCGGTAAATATAGTGTTCATCATTTTTCAGATTGTGAATTTTGGCTGCC 580
QY      161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnIysLeuCysAspGlyIleAsn 180
      |||
DB      581 TTGTATTAATCCTTTTGGGAATTTTAAACCCCACTTCAAAAACCTAGTGTATCAAC 640
QY      181 LysMetLeuAspGluGluAsnIle 188
      |||
DB      641 AAAATGTGATGAAGAACAATCA 664

RESULT 7
AX877673 1729 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 12578 from Patent EP1074617.
DEFINITION AX877673
ACCESSION AX877673.1 GI:40032409
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Ota,T., Isogai,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Makamatsu,A., Nagai,K. and Otsuki,T.
TITLE Primers for synthesising full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 12578 07-FEB-2001;
Research Association for Biotechnology (JRP)
FEATURES
source
1. 1729
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
71..667
/note="unamed protein product"
/codon_start=1
CDS

```

/protein_id="CAE98929.1"
/db_xref="GI:40032410"
/translation="MHSRASESKVATDVNSKNLAVOAKKILIGKMSKSIATLLID
DTSSVLDLRYTRVETYNKKKEAKKIKLIKTYIKALILRNQFNODELALMEKF
KKVHQIAMTVSFFHQVDYTRDRNLVSLNLECREMHLQIILQRHLTASHGKVVNVPD
HFSDEFLAALYNPPGNFKPHLOKLCIDGINKMLDEENT"

ORIGIN

Alignment Scores:

Pred. No.: 1,89e-75 Length: 1729
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x AX877673 (1-1729)

```

QY      1 MetAlaThrAspValaPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
      :::
DB      101 GGGCCACAGATGCTTTTAATTCAAAACCTGGCCGTTCAAGCACAAAAGAGATCTTG 160
QY      21 GlyLysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      |||
DB      161 GGTAAATGCTGTCCAAATCCATCCGACCACTTAAATAGACACAAAGTAGTAGG 220
QY      41 LeuAspGluLeuTyrArgValaThrArgIuTyrThrGlnAsnIysIleGluAlaGluLys 60
      |||
DB      221 CTGGATGAGCTTACAGAGTGAACGAGGAGTACACCAAAACAAAGAGGAGAGAG 280
QY      61 LysIleLysAsnLeuIleIysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
      |||
DB      281 ATCATCAAGAACTCATCAAGACAGTATCAAGCTGCCATCTTAAAGAAATTAATCAG 340
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysIleLysValaHisGlnLeuAla 100
      |||
DB      341 TTTAAATCAAGATGAGCTTGAATGAGGAAATTTAAAGAAAGATTCATCAGCTTGCT 400
QY      101 MetThrValaIleSerPheHisGlnValaAspTyrThrPheAspArgAsnValaLeuSerArg 120
      |||
DB      401 ATGACCGTGCTGATTCCATCAGTGATGATTATACCTTTGACCGGAATGTTATCCAG 460
QY      121 LeuLeuAsnGlnCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||
DB      461 CTGTAAATGAATGACAGAGATGCTGCACCAATATATTCAGGCCACCTCACTGCCAAG 520
QY      141 SerHisGlyArgValaAsnAsnValaPheAspHisPheSerAspCysGluPheLeuAlaAla 160
      |||
DB      521 TCACATGACGCGGTAAATATAGTGTTCATCATTTTTCAGATTGTGAATTTTGGCTGCC 580
QY      161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnIysLeuCysAspGlyIleAsn 180
      |||
DB      581 TTGTATTAATCCTTTTGGGAATTTTAAACCCCACTTCAAAAACCTAGTGTATCAAC 640
QY      181 LysMetLeuAspGluGluAsnIle 188
      |||
DB      641 AAAATGTGATGAAGAACAATCA 664

RESULT 8
AK001850 1729 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ10988 f1s, clone PLACB1001920, b1ghly s2mlar
DEFINITION to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK001850
VERSION AK001850.1 GI:7023373
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
Makamatsu,A., Hayaishi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,

```

Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Negahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuna, M., Shitatori, A., Sudo, H., Hosokiri, T., Kaku, Y., Kodaira, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsura, N., Saito, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakami, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirao, K., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoda, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi, F., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A., Sasaki, N., Aotaka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Komiyama, H., Saitoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Omori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
Natl. Genet. 36 (1), 40-45 (2004)
14702039

2
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Negahara, K., Masuno, Y., and Oshima, A. NEDD human cDNA sequencing project

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
Unpublished
3 (bases 1 to 1729)
Isogai, T. and Otsuki, T.
Direct Submission

COMMENT
Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@rhi.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDD human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan. cDNA full insert sequencing: Research Association for Biotechnology, cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES
source
Location/Qualifiers

1..1729
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACB1001920"
/issue_type="Placenta"
/clone_lib="PLACB1"
/note="Cloning vector: pME18SFL3"

ORIGIN

Alignment Scores:
Pred. No.: 1,898-75 Length: 1729
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x AK001850 (1-1729)

Qy 1 MetAlaThrAspValPheAsnSerIleValGlnAlaGlnLysIleLeu 20
Db 101 GTGGCCAGATGCTTTTAATTCAAAACCTGGCCGTCACGACAAAAGATCTTG 160
Qy 21 GLLYMeValSerIleValThrThrLeuIleAspThrSerSerGluVal 40
Db 161 GGTAAATGTGTCCAAATTCATTCGCCACCTTTATAGACACAAAGTAGAGGTG 220
Qy 41 LeuAspGluLeuValArgValThrArgIuValThrGlnAsnLysGluVal 60
Db 221 CTGATGAGCTTACAGAGTACAGAGGAGTACAGGAGTACCCAAAACAAAGAGGAGAG 280
Qy 61 LysIleLeuAsnLeuIleValThrValIleValLeuAlaIleLeuValArgAsnGln 80
Db 281 ATCATAGAAAGCTCATCAAGACGATCACTAGCTGCCATCTTTATAGAAATATCAG 340
Qy 81 PheAsnGlnAspGluLeuValLeuMetGluValPheLysLysValHisGlnLeuAla 100
Db 341 TTTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGTTCATCAGCTTGT 400
Qy 101 MetThrValValSerPheHisGlnValAspValThrPheAspArgAsnValLeuSerArg 120
Db 401 ATGACCGTGTGTCAGTTTCCATCAGTGGATTAATTAATCCTTTGACCGGAATGTTATCCAG 460
Qy 121 LeuLeuAsnGluCysArgGluLeuLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db 461 CTGTAAATGAATGACAGAGATGTCGACCAATATTCATTCAGGCGACCTCAGCCAG 520
Qy 141 SerHisGluValArgValAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
Db 521 TCACATGACGAGGATTAATATGTTGATCATTTTTCAGATTTGTAATTTTGGCTGCC 580
Qy 161 LeuValAspProPheGluValAsnPheLysProHisLeuGluLeuValLeuValAsn 180
Db 581 TTGTATTAATCCTTTTGGGATTTTAAACCCCACTTACAAAACCTATGATGATCAAC 640
Qy 181 LysMetLeuAspGluGluLeuIle 188
Db 641 AAAATGTTGATGAGAGAAACATTA 664

RESULT 9
LOCUS BC005352 1814 bp mRNA linear PRI 24-FEB-2004
DEFINITION Homo sapiens tumor necrosis factor, alpha-induced protein 8, mRNA (cDNA clone MGC:12451 IMAGE:3997650), complete cds.
ACCESSION BC005352
VERSION BC005352.1 GI:13529163
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1814)
Klausner, R.D., Collins, F.S., Wagner, L., Shemmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buecort, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, L., Marusina, K., Farmer, A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uebli, T.B., Toshiyuki, S., Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulik, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shivehenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, J., Smal, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

TITLE

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED
1247932

REFERENCE
2 (bases 1 to 1814)

AUTHORS
Strausberg, R.

TITLE
Direct Submission

JOURNAL
Submitted (27-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT
Contact: MGC help desk
Email: cga@bsf-research.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.sbc.stanford.edu>
Contact: (Dickson, Mark) mcdonald@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES
source

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAL Plate: 16 Row: b Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657123.
Location/Qualifiers
1..1814
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12451 IMAGE:3997650"
/tissue_type="Bladder, carcinoma"
/clone_id="N1H MGC_53"
/lab_host="DH10B"
/note="vector: pDNR-LIB"

Oy		21	GlyuysmetValSerlysSerIlealaThrThreuleuLeuaspThrserSerGluVal	40
Dd		234	GgTAAATGGGTCCAAATCCATGCCACCACTTAATAGCGACAAGTAGTGAGGTG	293
Oy		41	LeuAspGluLeuTyraArgValThrArxGluTyrThrGlnasnLysGluValagLylus	60
Dd		294	CTGAGCAGGCTCTACAGGTGCACCAGGAGTACCCAAAACAAGAAGAGGACAGAG	353
Oy		61	LysIlleLysasnLeuIleLysThrValIlleLysLeuAlalleuTyraGdnasnGln	80
Dd		354	ATCATCAAAGAACCTTCACACAGACAGCATCAAGCTGGCCATTCTTTAAGGAATAATCAC	413
Oy		81	PheasnGlnasprGluLeuAlaleMetGlnLysPheLysLysValHisGlnLeuAla	100
Dd		414	TTTAATCAAGATGAGCTTACATGTGATGAGAAATTTAAGAAGAAAGTTCAATCAGCTTCT	473
Oy		101	MetThrValValSerPheHisGlnValAspTyrThrPheAsparGdnValLeuSerArx	120
Dd		474	ATGACCGGTGCAGTTTCCATCAGGTGAGTATATACCTTTGACCGGAATGTATTATCCAG	533
Oy		121	LeuLeuasnGluCyaaArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys	140
Dd		534	CTGTAAATGAAATGCACAGAGATGCTGCACCAATCTTAAGCGCCACTCATCGCCAAG	593
Oy		141	SerIsigLyarGValaasnValPheAspHisPheSerAspCysGluPheLeuAlaAla	160
Dd		594	TCAATATGACGGGTAAATTAATGCTTTGATCATTTTTCAAGATTGTGAATTTTGGCTGCC	653
Oy		161	LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleasn	180
Dd		654	TTGTATATATCTTTTGGGAATTTTAAACCCCACTTAACAATACTATGTGATGATCAAC	713
Oy		181	LysMetLeuAspGluGluAsnIle	188
Dd		714	AAAAATGTTGATGAAAGAAACATA	737
RESULT 10				
AF099935				
LOCUS		AF099935	1915 bp mRNA linear PRI 12-NOV-1998	
DEFINITION		Homo sapiens MDC-3.13 isoform 2 mRNA, complete cds.		
ACCESSION		AF099935		
VERSION		AF099935.1	GI:3860092	
KEYWORDS		.		
SOURCE		Homo sapiens (human)		
ORGANISM		Homo sapiens (human)		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE		1 (bases 1 to 1915)		
AUTHORS		Nietfeld,W. and Meyerhans,A.F.		
TITLE		Identification of cellular factors involved in the differentiation		
JOURNAL		of dendritic cells		
REFERENCE		Unpublished		
AUTHORS		2 (bases 1 to 1915)		
TITLE		Nietfeld,W. and Meyerhans,A.F.		
JOURNAL		Direct Submission		
FEATURES		Submitted (13-OCT-1998) Max Planck Institute for Molecular		
		Genetics, Imneestrase 73, Berlin 14195, Germany		
source		Location/Qualifiers		
		1..1915		
		/organism="Homo sapiens"		
		/mol_type="mRNA"		
		/db_xref="taxon:9606"		
		/cell_type="adherent monocytes treated with GM-CSF and		
		IL-4"		
CDS		85..681		
		/note="MDC-3.13/2"		
		/codon_start=1		
		/product="MDC-3.13 isoform 2"		
		/protein_id="AAC72975.1"		
		/db_xref="GI:3860093"		
		/translation="MHSEAESEKSVATDVFNASKLAIVQAKIKLGKMSKIATLTLD		
		DTSEVDIELYRTVRTEYONKRAEKIKNLITVTKIALLYRNNOFNDELIMKEKF		

ORIGIN
KKKVVHOLAMTVSFEHQVDYTFEDRNVL SRLNCEEMLHOI IORHLAKSHGRVNVNF
HFSDCEFIAMALVYVPGNFKXPHLQKLDGINKMLDEENI"

Alignment Scores:

Pred. No.: 2,1e-75 Length: 1915
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 98.75% Indels: 0
DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x AF099935 (1-1915)

QY 1 MetCAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
115 GTGGCCACAGATGCTTAAATCCAAAACCTGGCCGTCAGGCAAAAGAGATCTTG 174
QY 21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
175 GGTAAATGGTGTCCAAATCCATCGCACCACTTAATAGACGACAAAGTAGAGGTG 234
QY 41 LeuAspGluLeuTyrArgValThrArgLysThrGlnAsnLysLysGluAlaGluLys 60
235 CTGGATGAGCTTACAGAGTACAGGAGGAGTACCCAAACAAAGAGGAGGAGAG 234
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
295 ATCATCAAGAACTCATCAAGACAGATCATAGTGGCCATCTTTATAGAAATATCAG 354
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
355 TTTAAATCAAGATGAGTATGATGAGAGAAATTTAAGAAAGTTCATCAGCTTGT 414
QY 101 MetThrValIleSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
415 ATGACCGTGTCTGATTCATCAGTGGATTTATACCTTTACCGGAATGTATTACAGG 474
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
475 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCAGTCCAG 534
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
535 TCACATGAGACGGGTAAATATGATGATGATTTTCAATTTGTGAATTTTGGCTGCC 594
QY 161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
595 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATATCAAC 654
QY 181 LysMetLeuAspGluGluAsnIle 188
655 AAAATGTGGATGAGAGAACATTA 678
Db

RESULT 11
CQ726075 1988 bp DNA linear PAT 03-FEB-2004
LOCUS CQ726075
DEFINITION Sequence 12009 from Patent WO02068579.
ACCESSION CQ726075
VERSION CQ726075.1 GI:42288070
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE kites, such as nucleic acid arrays, comprising a majority of
humaneurons or transcripts, for detecting expression and other uses
therof
JOURNAL Patent: WO 02068579-A 12009 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES location/Qualifiers

source 1. .1988

/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 2,18e-75 Length: 1988
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 0
Query Match: 98.75% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x CQ726075 (1-1988)

QY 1 MetCAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
204 GTGGCCACAGATGCTTAAATCCAAAACCTGGCCGTCAGGCAAAAGAGATCTTG 263
QY 21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
264 GGTAAATGGTGTCCAAATCCATCGCACCACTTAATAGACGACAAAGTAGAGGTG 323
QY 41 LeuAspGluLeuTyrArgValThrArgLysThrGlnAsnLysLysGluAlaGluLys 60
324 CTGGATGAGCTTACAGAGTACAGGAGGAGTACCCAAACAAAGAGGAGGAGAG 383
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
384 ATCATCAAGAACTCATCAAGACAGATCATAGTGGCCATCTTTATAGAAATATCAG 443
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
444 TTTAAATCAAGATGAGTATGATGAGAGAAATTTAAGAAAGTTCATCAGCTTGT 503
QY 101 MetThrValIleSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
504 ATGACCGTGTCTGATTCATCAGTGGATTTATACCTTTGACCGGAATGTATTACAGG 563
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
564 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACTCAGTCCAG 623
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
624 TCACATGAGACGGGTAAATATGATGATGATTTTCAATTTGTGAATTTTGGCTGCC 683
QY 161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
684 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATATCAAC 743
QY 181 LysMetLeuAspGluGluAsnIle 188
744 AAAATGTGGATGAGAGAACATTA 767
Db

RESULT 12
AF099936 2003 bp mRNA linear PRI 12-NOV-1998
LOCUS AF099936
DEFINITION Homo sapiens MDC-3.13 isoform 1 mRNA, complete cds.
ACCESSION AF099936
VERSION AF099936.1 GI:3860094
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Niefeld, W. and Meyerhans, A.F.
TITLE Identification of cellular factors involved in the differentiation
of dendritic cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2003)

AUTHORS Nietfeld, W. and Meyerhans, A.F.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1998) Max Planck Institute for Molecular
Genetics, Ihmestrasse 73, Berlin 14195, Germany
FEATURES Location/Qualifiers
source 1. .2003

CDS

Alignment Scores:	
Pred. No.:	2, 2e-75
Score:	950.00
Percent Similarity:	99.47%
Best Local Similarity:	98.94%
Query Match:	98.75%
DB:	9
Gaps:	0
Length:	2003
Matches:	186
Conservative:	1
Mismatch:	1
Indels:	0
Gaps:	0

US-10-627-571-2 (1-188) x AF099936 (1-2003)

OY	1	MetaIatrraapValPheasnsertlysasntleuAlaValGlnlAglnlylsylleu	20
		
Db	204	GTGGCCACGAGTCTTTATATTCAAAAACCTGGCCGTTCCAGCACAAAAAGAACTTGG	263
OY	21	GlylvsMetValSerlyseSerlleaIatrrhLeuIlasaprrhSerSerGluVal	40
Db	264	GGTAAAAATGGTCTCCAAATCCATCCGCACACCTTATATAGACACACAGTATGAGGTG	323
OY	41	LeuAapGluLeuTyraArgValThraArgIuTyrrhGlnasnlYalysGluAlaGluys	60
Db	324	CTGATGTAGCTCTTACAGAGTGAACAGGAGATCACCCAAAAACAAGAGGCGACAGAGAG	383
OY	61	LysIlelysaAsnleuIlelysrhValIlelyslleuAlaIlleuTyraArgaAsnGln	80
Db	384	ATCATCAAGAAACCTTATCATGAACAGTCATCAAGCTGGCCATTTTATATAGAAATATTCAG	443
OY	81	PheaEngInaapGluLeuAlaLeuMetGluysPheLyelYalysValHsGlnleuAla	100
Db	444	TTTATATCAAGATGAGCTTACATTGATGAGAAATTTTAAGAGAAAGTTCTACAGCTTGCT	503
OY	101	MetThrValValSerPheHsGlnValaAspTyrrhPheAspArgaAsnValleuSerArg	120
Db	504	ATGACCGGTGTGATTTCCATCAGGGTGATTAATCACTTTGACCGGAAATGTTATCCAGG	563
OY	121	LeuIleuAsnGluCyArgGluMetleuHsGlnIlleGlnIargHsleuThrAlaLys	140
Db	564	CTGTTTAAATGAATGACGAGAGATGCTGCACCAATATTCAGGCCACCTCAGTGCACAG	623
OY	141	SerHsGlyArgValAsnaAsnValPheAspHsAspSerAspCyGluPheleuAlaAla	160
Db	624	TCACATGGACGGGTATATATGATGTTGATCATTTTTCAGATTTGTAATTTTGGCTGCC	683
OY	161	LeuTyraAsprrhPheGlyAsnPheLysProHsleuGlnLysleuCyAspGlyIleAsn	180
Db	684	TTGTATATATCCTTTGGGGAATTTTAAACCCCACTTACAAAAATATGTATGTATCAAC	743
OY	181	LysMetLeuAapGluGlnAsnIle	188
Db	744	AAATATGTTGATGAGAGAACATA	767

RESULT	13
LOCUS	BC007014
DEFINITION	Homo sapiens tumor necrosis factor, alpha-induced protein 8, mRNA
ACCESSION	BC007014
VERSION	BC007014.1
KEYWORDS	MG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

TITLE	JOURNAL	PUBLISHED	REFERENCE	AUTHORS	TITLE	JOURNAL
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	Proc. Natl. Acad. Sci. U.S.A.	99 (26), 16899-16903 (2002)				
2 (bases 1 to 2081)		12477932				
Strausberg, R.						
Submitted (30-Apr-2001)	National Institutes of Health, Mammalian					

FEATURES	Location/Qualifiers
SOURCE	1. .2081

gene

QY 121 LeuLeuAaNgLuCyArGluMeLeuHieGlnIleIeGlnArGHisLeuThraIaLys 140
 Db 65408 CTGTTAAATGAATGCGAGAGATGCTGCACCAATATTCAGGCCACCTCCTAGCCAAAG 65467
 QY 141 SerHisGlyArGValaAsnAnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
 Db 65468 TCACATGACGAGGCTTAATATATGTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 65527
 QY 161 LeuTyraAspProPheGlyAsnPhelyProHisLeuGlnIlyLeuCyAspGlyIleAsn 180
 Db 65528 TTGTAAATATCCCTTTGGGAATTTTAAACCCCACTTCAAAACTATGATGTATCAAC 65587
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 65588 AAAATGTTGATGAGAGAAACATA 65611

RESULT 15
 AC026795/c 156277 bp DNA linear PRI 21-SEP-2001
 LOCUS Homo sapiens chromosome 5 clone CTD-2287122, complete sequence.
 DEFINITION AC026795
 AC026795.6 GI:15718543
 VERSION
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 TITLE DOE Joint Genome Institute and Stanford Human Genome Center.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 156277)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 3 (bases 1 to 156277)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (31-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 REFERENCE 4 (bases 1 to 156277)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT On Sep 21, 2001 this sequence version replaced gi:15383783.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 99.5% of Sequence;
 Estimated Total Number of Errors is 0.4.
 STS Content:
 SHGC-7119 G14211.
 Location/Qualifiers
 1..156277
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CTD-2287122"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,96e-73 Length: 156277
 Score: 950.00 Matches: 186
 Percent Similarity: 99.47% Conservative: 1
 Best Local Similarity: 98.94% Mismatches: 1
 Query Match: 98.75% Indels: 0
 DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x AC026795 (1-156277)

QY 1 MetAlaThrAspValPheAsnSerIyAsnLeuAlaValGlnAlaGlnIlySerIleLeu 20
 Db 128234 GTGGCCACAGATGCTTTTAATTCAAAACTGGCCGTTCAAGGCACAAAGAAAGATCTTG 128175
 QY 21 GlyLysMetValSerIySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 128174 GGTAAATATGTCCTCAATATCCATGCCACACCTTAATAGACACAAAGTATGTAGGTG 128115
 QY 41 LeuAspGluLeuTyraGValThrArgIuTyThrGlnAsnIlyLeuAlaGluLys 60
 Db 128114 CTGATGAGCTCTACAGAGTGACAGGAGATCACCCAAACAAAGAGAGCAAG 128055
 QY 61 LysIleLysAsnLeuIleIyThrValIleLysLeuAlaIleLeuTyraArgAsnGln 80
 Db 128054 ATCATCAAGAACCTCATCAAGCACTCATCAAGCTGGCCATCTTTATAGAAATATACG 127995
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 127994 TTTAATCAAGATGAGCTAGCATTGATGAGAAATTTAAGAAAGAACTTCATCAGCTTGT 127935
 QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
 Db 127934 ATGACCGGTGTCAAGTTCATCAGGTGATTAATACCTTGGACCGGAATGTGTTATCCAGG 127875
 QY 121 LeuLeuAaNgLuCyArGluMeLeuHieGlnIleIeGlnArGHisLeuThraIaLys 140
 Db 127874 CTGTTAAATGAATGAGAGAGATGCTGCACCAATATTCAGCGCCACTCCTCAGCAG 127815
 QY 141 SerHisGlyArGValaAsnAnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
 Db 127814 TCACATGACGAGGCTTAATATATGTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 127755
 QY 161 LeuTyraAspProPheGlyAsnPhelyProHisLeuGlnIlyLeuCyAspGlyIleAsn 180
 Db 127754 TTGTAAATATCCCTTTGGGAATTTTAAACCCCACTTCAAAACTATGATGTATCAAC 127695
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 127694 AAAATGTTGATGAGAGAAACATA 127671

RESULT 16
 AC138612/c 158057 bp DNA linear PRI 30-MAY-2003
 LOCUS Homo sapiens BAC clone CTD-2233E22 from 2, complete sequence.
 DEFINITION AC138612
 AC138612.1 GI:27597054
 ACCESSION
 VERSION AC138612.1 GI:27597054
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 JOURNAL Unpublished (2001)
 REFERENCE 2 (bases 1 to 158057)
 AUTHORS Sultston,J.E. and Wilson,R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 3 (bases 1 to 158057)
 AUTHORS Tomlinson,C. and Bielicki,L.
 TITLE The sequence of Homo sapiens BAC clone CTD-2233E22
 JOURNAL Unpublished (2001)
 REFERENCE 4 (bases 1 to 158057)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (11-JAN-2003) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 5 (bases 1 to 158057)
 AUTHORS Wilson,R.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAY-2003) Department of Genetics, Washington

COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

----- Genome Center
Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu

----- Summary Statistics

Center project name: H_MS2293E22

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTD-2293E22 is from a release of the human BAC library CTD. The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
VECTOR: pBelorBAC11
Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source Location/Qualifiers

```

1..158057
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="2"
  /map="2"
  /clone="CTD-2293E22"
  /clone_1b="CTD"
repeat_region      4..83
                    /rpt_family="MIR"
repeat_region      300..719
                    /rpt_family="ERV1"
repeat_region      1321..1459
                    /rpt_family="L1"
repeat_region      1518..1547
                    /rpt_family="AT_rich"
repeat_region      1964..1996
                    /rpt_family="(TA)n"
repeat_region      2532..2575
                    /rpt_family="A-rich"
repeat_region      2639..2841
                    /rpt_family="L2"
repeat_region      2859..2886
                    /rpt_family="AT_rich"
repeat_region      3179..3201
                    /rpt_family="AT_rich"
repeat_region      4351..4541
                    /rpt_family="MIR"
repeat_region      5128..5766
                    /rpt_family="L1"
repeat_region      5767..6328
                    /rpt_family="L1"

```

```

repeat_region      6329..7499
                    /rpt_family="L1"
repeat_region      7503..7872
                    /rpt_family="L1"
repeat_region      7920..8106
                    /rpt_family="MIR"
repeat_region      8621..8849
                    /rpt_family="ERV1"
repeat_region      9113..9415
                    /rpt_family="Alu"
repeat_region      12251..12442
                    /rpt_family="L1"
repeat_region      14323..14361
                    /rpt_family="AT_rich"
repeat_region      14630..14899
                    /rpt_family="Alu"
repeat_region      15273..15541
                    /rpt_family="L1"
repeat_region      15534..17744
                    /rpt_family="L1"
repeat_region      18238..18436
                    /rpt_family="MIR2_type"
repeat_region      18725..18973
                    /rpt_family="MIR2_type"
repeat_region      19548..19664
                    /rpt_family="L2"
repeat_region      19666..19723
                    /rpt_family="(TAAG)n"
repeat_region      20144..20192
                    /rpt_family="polypurine"
repeat_region      20415..20468
                    /rpt_family="AT_rich"
repeat_region      20476..20771
                    /rpt_family="Alu"
repeat_region      21591..22250
                    /rpt_family="L1"
repeat_region      22475..24483
                    /rpt_family="L1"
repeat_region      24487..24515
                    /rpt_family="AT_rich"
repeat_region      25137..25196
                    /rpt_family="(CA)n"
repeat_region      26530..26604
                    /rpt_family="L2"
repeat_region      27433..28325
                    /rpt_family="L2"
repeat_region      28361..28469
                    /rpt_family="MIR"
repeat_region      28603..28653
                    /rpt_family="(CATATA)n"
repeat_region      29092..29113
                    /rpt_family="AT_rich"
repeat_region      29379..29680
                    /rpt_family="Alu"
repeat_region      29871..30008
                    /rpt_family="MIR1_type"
misc_feature       30569..30609
                    /note="Sequence derived from PCR product of project DNA."
repeat_region      31531..31588
                    /rpt_family="GA-rich"
repeat_region      31939..32107
                    /rpt_family="ERV1"
repeat_region      32426..32605
                    /rpt_family="Alu"
repeat_region      32814..32878
                    /rpt_family="MIR"
repeat_region      33458..33570
                    /rpt_family="L2"
repeat_region      33887..33984
                    /rpt_family="(GAAA)n"
repeat_region      35523..35595
                    /rpt_family="AT_rich"
repeat_region      36412..36602

```

```

repeat_region      /rpt_family="MIR"
                    37311..37413
repeat_region      /rpt_family="L1"
                    37414..37724
repeat_region      /rpt_family="Alu"
                    37725..37879
repeat_region      /rpt_family="L1"
                    38402..38429
repeat_region      /rpt_family="AT_rich"
                    38459..38494
repeat_region      /rpt_family="AT_rich"
                    39513..39548
repeat_region      /rpt_family="CANA-Glu-GAG"
                    39654..39929
repeat_region      /rpt_family="L1"
                    39956..40148
repeat_region      /rpt_family="MIR"
                    41023..41048
repeat_region      /rpt_family="(A)n"
                    41243..41431
repeat_region      /rpt_family="L1"
                    41475..41561
repeat_region      /rpt_family="L1"
                    41789..41811
repeat_region      /rpt_family="AT_rich"

```

Alignment Scores:

```

Pred. No.:      1,996-73      Length:      158057
Score:          950.00      Matches:      186
Percent Similarity: 99.47%      Conservative: 1
Best Local Similarity: 98.94%      Mismatches: 1
Query Match:     98.75%      Indels:      0
DB:              9      Gaps:      0

```

US-10-627-571-2 (1-188) x AC138612 (1-158057)

```

QY      1 MetcAtThrAspValPheAsnSerIysAsnLeuAlaValGlnAglInySylIleLeu 20
DB      135468 GTGGCCACAGATGCTTTAATTCACAAAACCTGGCGCTCAGCACAAAAGAGATCTTG 135409
QY      21 GILySMeCValISerIySerIleAlaThrIleuIleAspAspThrSerSerGluVal 40
DB      135408 GGTAAATGCTGTCCAATCCAGCGCACCTTAATAGACACAAAGTGTGAGTGC 135349
QY      41 LeuAspGluLeuTyrrArGValThirArGluTyrrThGlnAsnIySglInySylVal 60
DB      135348 CTGGATGAGCTCTACAGAGTGCAGGAGTACACCCAAAACAAAGAGGCGCAGAG 135289
QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArGAsnAngIn 80
DB      135288 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGGAATATCAG 135229
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysIySylValIhIsgInLeuAla 100
DB      135228 TTTAAATCAAGATGAGCTAGATTCATGAGAAATTTAAGAAAGTTCATCAGCTTGC 135169
QY      101 MetThrValValSerPheHISGInValAspTyrrThPheAspArgAsnValLeuSerArg 120
DB      135168 ATGACCGTGCTCAGTTTCCATCCAGTGGAGTATTAACCTTTGACCGGAATGTTATCCAG 135109
QY      121 LeuLeuAsnGluLysArGluMetLeuHISGInIleIsgInArgHISLeuThrAlaLys 140
DB      135108 CTGTTAAATGATGACGAGAGATGCTGCACCAATCATTCAGGCCACCTCATGCGCAAG 135049
QY      141 SerHISGILyArGValAsnAsnValPheAspHISPheSerAspCysGluPheLeuAlaLys 160
DB      135048 TCACATGAGACGGGTATATATATGTTGATCATTTTTCAGATGTGAATTTTGGCTGCC 134989
QY      161 LeuTyrrAsnProPheGlyAsnPheLysProHISLeuGlnIyLeuLysAspGlyIleAsn 180
DB      134988 TTGATTAATTCCTTTTGGAAATTTTAAACCCCACTTACAAAACATATGATGATCAAC 134929
QY      181 LysMetLeuAspGluGlnAsnIle 188

```

DB 134928 AAAATGTCGATGAGAGACACTA 134905

```

RESULT 17
BD135356
LOCUS      BD135356      1943 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION 110 human secretory proteins.
ACCESSION  BD135356
VERSION    BD135356.1 GI:23230301
KEYWORDS   JP 2002508167-A/107.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```

```

REFERENCE
AUTHORS    Moore,P.A., Ruden,S.M., Carter,K.C., Shi,Y., Rosen,C.A.,
            Soppet,D.R., Caou,H., Wei,Y.F., Florence,K., Duan,R.D.,
            Florence,C., Greene,J.M., Feng,P., Ferrite,A.M., Yu,G.L., Janat,F.
            and Ni,J.

```

```

TITLE
JOURNAL
COMMENT
HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002508167-A/107
PD 19-MAR-2002

```

```

PR 17-DEC-1998 JP 2000539040
PR 18-DEC-1997 US 60/070 923,18-DEC-1997 US 60/068 007 PR
18-DEC-1997 US 60/068 057,18-DEC-1997 US 60/068 006 PR
18-DEC-1997 US 60/068 008,18-DEC-1997 US 60/068 054 PR
18-DEC-1997 US 60/068 064,18-DEC-1997 US 60/068 053 PR
19-DEC-1997 US 60/068 169,19-DEC-1997 US 60/068 368 PR
19-DEC-1997 US 60/068 367,19-DEC-1997 US 60/068 369 PR
19-DEC-1997 US 60/068 365

```

```

PI PAUL A MOORE,STEVEN M RUBEN,KENNETH C CARTER,YANGU SHI,CRAIG
PI A ROSEN,
PI DANIEL R SOPPET,HARA CAOU,YING FEI WEI,KIMBERLY FLORENCE, PI
PI ROSANNE D DUAN,
PI CHARLES FLORENCE,JOHN M GREENE,PING FENG,ANN M FERRIE,GUO PI
LIANG YU

```

```

PI FORD JANAT,JIAN NI
PC C12N15/09,A61K38/00,A61K48/00,A61P9/00,A61P9/10,A61P15/00, PC
A61P25/00,
PC A61P25/02,A61P25/14,A61P25/16,A61P25/18,A61P25/22,A61P25/24,
PC A61P25/28,

```

```

PC A61P29/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/
10,
PC C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53,C12N15/00, PC
A61K37/02,
PC C12N5/00

```

```

CC n equals a,t,g,or c
CC n equals a,t,g,or c
CC n equals a,t,g,or c
CC n equals a,t,g,or c
CC Key
FH Key
FT source

```

```

FEATURES
source location/Qualifiers
1..1943
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      3,216-75      Length:      1943
Score:          948.00      Matches:      186
Percent Similarity: 98.94%      Conservative: 0
Best Local Similarity: 98.94%      Mismatches: 2
Query Match:     98.54%      Indels:      0
DB:              6      Gaps:      0

```

US-10-627-571-2 (1-188) x BD135356 (1-1943)

```

QY      1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAglInlySylleu 20
DB      120 ATGGCCACGAGATCTTTAATTCGAAAAACCTGCGCTTANCCACAAAAAGAGATCTTG 179
QY      21 GlylMetValSerIysSerIleAlaThrThleuIleAspAspThrSerSerGluVal 40
DB      180 GGTAAATGGTGTCCAAATTCATCGCCACCACTTATAGACACACAGATAGTGGGTG 239
QY      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnIysIysGluVal 60
DB      240 CTGGATGAGCTTACAGAGTACGACGAGGAGTATACACCAAAACAAAGAGGAGGAG 299
QY      61 LysIleLeuAsnLeuIleTyThrValIleIysLeuAlaIleLeuTyArgAsnGln 80
DB      300 ATCATAGAGAACTCATCAAGACGATCAACCTGCGCCCTTTATAGAAATTAATCAG 359
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLeuPheIysIysValIhIsglnLeuAla 100
DB      360 TTTAAATCAAGATAGCTAGCTAGTGAAGAGAAATTTAAGAAAGATTTCATCAGCTTGT 419
QY      101 MetThrValIserPhehIsglnValAspTyThrPheAspArgAsnValLeuSerArg 120
DB      420 ATGACCGTGGTCAAGTTCCATCAGGTGATTAATACCTTGACCGGAATGTTATCCAG 479
QY      121 LeuLeuAsnGluCySarGluIleMetLeuHIsGlnIleIleGlnArgHIsleuThrAla 140
DB      480 CTGTTAAATGAATGACGAGAGATGCTGCACCAATCATTCAGCGCCACCTCCTCCAG 539
QY      141 SerHISGlyArgValAsnAsnValPheAspHISpSerAspGluPheLeuAlaIa 160
DB      540 TCAACAGAGAGGGTTAAATATGTTGATCATTTTCAGATTGTGAATTTTGGCTGCC 599
QY      161 LeuTyArgAsnProPheGlyAsnPhelysProHISleuGlnIysLeuCySarGlyIleAsn 180
DB      600 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTCAAAATAATATGTATGTTATCAC 659
QY      181 LysMetLeuAspGluGluAsnIle 188
DB      660 AAAATGTTGATGACAGAGAACATTA 683

```

RESULT 18
AK097284
LOCUS AK097284 2502 bp mRNA linear PRI 30-JAN-2004
DEFINITION Homo sapiens cDNA FLJ39965 f1s, clone SPLEN2027157, highly similar to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK097284
VERSION AK097284.1 GI:21756982
KEYWORDS oligo capping, f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakematsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohtsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, Y., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, T., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Hosoda, M., Hoota, T., Kusano, J., Kanehori, K., Takahashi, P., Hara, R., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuma, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,

Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hiki, J., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
 Nat. Genet. 36 (1), 40-45 (2004)
 14702039

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hiraoka, S., Tanai, H., Watanabe, S., Ishida, S., Ono, Y., Hoota, T., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakematsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi, Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 unpublished
 3 (bases 1 to 2502)
REFERENCE Isogai, T. and Yamamoto, J.
AUTHORS Direct Submission
TITLE Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kametani, Kiserazu, Chiba 292-0812, Japan
JOURNAL (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
 Location/Qualifiers
 1..2502
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="SPLEN2027157"
 /tissue_type="spleen"
 /clone_lib="SPLEN2"
 /note="Cloning vector: pME18FL3"

FEATURES
 source

ORIGIN
 Alignment Scores:
 Pred. No.: 9,486-75 Length: 2502
 Score: 944.00 Matches: 185
 Percent Similarity: 99.478 Conservative: 2
 Best Local Similarity: 98.408 Mismatches: 1
 Query Match: 98.134 Indels: 0
 DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x AK097284 (1-2502)

```

QY      1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAglInlySylleu 20
DB      120 ATGGCCACGAGATCTTTAATTCGAAAAACCTGCGCTTANCCACAAAAAGAGATCTTG 179
QY      21 GlylMetValSerIysSerIleAlaThrThleuIleAspAspThrSerSerGluVal 40
DB      180 GGTAAATGGTGTCCAAATTCATCGCCACCACTTATAGACACACAGATAGTGGGTG 239
QY      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnIysIysGluVal 60
DB      240 CTGGATGAGCTTACAGAGTACGACGAGGAGTATACACCAAAACAAAGAGGAGGAG 299
QY      61 LysIleLeuAsnLeuIleTyThrValIleIysLeuAlaIleLeuTyArgAsnGln 80

```

```

Db      899 ATCATCAAAACCTCATCAAGACAGCATCAAGCTGCGCATCTTTATAGAAATATACG 958
Qy      81 Pheasnglaspgluleu1aleuwtg1ulyspheylalyeValh1sg1ueu1a 100
Db      959 TTTAATCAAGATGAGCTACATTTGAAGGAGAAATTTAAGAGAAAGTTATCATCGTTGCT 1018
Qy      101 MettThrValValSerPheH1sg1nvalasp1yrrThrPheaspArganValleuSerArg 120
Db      1019 ATGACCGGTGTCAGTTTCCATCATCAGGTGAGATTATACCTTTGACCGGAAATGTTATCCAG 1078
Qy      121 Leuleuasnglucyarsg1ueltueuH1sg1nle1le1gnarh1s1deurThralaly 140
Db      1079 CTGTTAAATGAATGACGAGAGATGCTGTACAAATCATTCAGGCGCACTCTGCGCAAG 1138
Qy      141 SerH1sg1yArValasnanValPheasPH1sPheSeraspCygluPheleu1a1a 160
Db      1139 TCACATGACGCGGTATATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 1198
Qy      161 LeuTyrrAspProPheglaspPhelysPheProH1sleuG1ulyleuCyaspG1ytlea 180
Db      1199 TTGATTAATCCTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTATGTATCAAC 1258
Qy      181 LysMetleuaspg1u1a1a1e 188
Db      1259 AAAATGTTGATGAAGAGAAACATA 1282

RESULT 19
LOCUS   BC009090              1728 bp      mRNA      linear      ROD 30-JUN-2004
DEFINITION Mus musculus tumor necrosis factor, alpha-induced protein 8, mRNA
            (cDNA clone MGC:11714 IMAGE:3965693), complete cds.
ACCESSION BC009090
VERSION   .
KEYWORDS  MGC.
SOURCE    Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1728)
AUTHORS   Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,
            Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
            Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
            Hopkins,R.P., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
            Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
            Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
            Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
            Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
            Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
            McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
            Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
            Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
            Fahey,U., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
            Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
            Bouffard,G.G., Blakesley,R.W., Touchman,D.W., Green,E.D.,
            Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
            Butlerfield,Y.S., Krzywinski,M.I., Skalska,U., Smallue,D.B.,
            Schnerich,A., Schein,J.E., Jones,S.J., and Marra,M.A.
            Generation and initial analysis of more than 15,000 full-length
            human and mouse cDNA sequences
            Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL   PUMED
PUBMED    12477932
REFERENCE 2 (bases 1 to 1728)
AUTHORS   Strausberg,R.
TITLE     Direct Submission
JOURNAL   Submitted (05-JUN-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT    Contact: MGC help desk
            Email: cgaabs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.

```

```

cDNA Library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louised, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAK Place: 17 Row: n Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19527361.
Location/Qualifiers
1..1728
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:11714 IMAGE:3965693"
/tissue_type="Mammary tumor, Metallothionien-TGF alpha
model, 10 month old virgin mouse. Taken by biopsy."
/clone_lib="NCI CGAP_Mam1"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
1..1728
/gene="Tnfalp8"
/notes="synonyms: Nded, Gq2-1, Ssc-2"
/db_xref="locusid:106869"
/db_xref="MGI:2147191"
70..666
/gene="Tnfalp8"
/codon_start=1
/product="TNF-induced protein"
/protein_id="AAH09090.1"
/db_xref="GI:14318586"
/db_xref="MGI:106869"
/db_xref="MGI:2147191"
/translacion="M1S2A3E3R3VATDVFNSKN1A4OAKKILGKMSKSIATTLID
DTSEVLELVKTVTSFQVETFTDRNVLSTLNECRELHEI1QRIHTAKSHGRVNNVFP
KKKXQMLTAVTSFQVETFTDRNVLSTLNECRELHEI1QRIHTAKSHGRVNNVFP
HFSDDFLAALYNPRGKRPKPHLQKLCIDGINKMBENI"

ORIGIN
Alignment Scores:
Pred. No.: 8,966-73 Length: 1728
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 10 Gaps: 0

US-10-627-571-2 (1-188) x BC009090 (1-1728)
Qy      1 MetAlaThrAspValPheasnserylsanleu1aValG1nAlG1nlyls1leu 20
Db      100 GTGGCTACAGATGCTTCAATCCAAAACCTGCGCTTCAGGCAAAAAGATCCCG 159
Qy      21 G1yLysMetleuasgluleu1aleuwtg1ulyspheylalyeValh1sg1ueu1a 40
Db      160 GGCAAAATGATGCCAAATCCATCGCCACACAGCTGATCGACACCAAGCAAGAGGTG 219
Qy      41 LeuAspGluLeuTyrrArgValThrArgGluTyrrThrGlnAsnLyseGlu1aGluys 60
Db      220 CTAGATGAGCTTACAGGTTGACAGAGATGTCACCCAGAACAGAGAGCGGAGAGG 279
Qy      61 Lys1lelyeAsnleu1leyserThrVal1leylsleu1a1leuTyrrArgan1ng 80
Db      280 GTCATCAAGAACTCATCAAGAGCGGTATCAAGCTGCGCTCTCCACAGAAATATCAG 339

```



```

TITLE
JOURNAL
REFERENCE
AUTHORS
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kelle,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., MacDonald,P., Major,J., Manning,T., Mathew,C.,
McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Miska,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunhphang,P., Pierre,N., Rachupka,A., Ramsamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schnupack,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J.,
Tophan,K., Travers,M., Vassiliou,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (32-JUN-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 184327)
Birren,B., Nusbaum,C., Lander,E., Abouelella,A., Allen,N.,
Anderson,M., Anderson,S., Archchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Bookhagalter,B., Camarata,J., Chang,U.,
Choopey,Y., Collamore,A., Cook,A., Cooke,P., Corum,B.,
D'Arnellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johanson,R., Jones,C., Kamat,A., Karatas,A., Kelle,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Maclean,C., MacDonald,P., Major,J., Manning,J., Mathew,C.,
McCarthy,M., Meldrim,J., Menus,L., Mihova,T., Mienga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunhphang,P., Pierre,N., Rachupka,A., Ramsamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schnupack,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J.,
Tophan,K., Travers,M., Vassiliou,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (05-AUG-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 5, 2004 this sequence version information replaced g1:49035085.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@ Broad.mit.edu
----- Project Information
Center project name: L21053
Center Clone name: 200_D_2
-----
FEATURES
source
location/Qualifiers
1..184327
/organism="Mus musculus"
/mol_type="genomic DNA"
/seq_xref="taxon:10090"
/chromosome="18"
/map="18"
/clone="RP24-200D2"
/clone_1b="RP24-200D2 Male Mouse BAC"
1..4
/note="clone boundary"
clone_end:SP6
site:MboI"
repeat_region
984..1007
/rpt_family="AT_rich"
/rpt_region
3199..3484

```

```
/rpt_family="B4A"
repeat_region      3506..3534
                    /rpt_family="CA)n"
repeat_region      complement(4228..4578)
                    /rpt_family="WTE"
repeat_region      4848..5049
                    /rpt_family="B3"
repeat_region      5526..5572
                    /rpt_family="AT_rich"
repeat_region      6521..6815
                    /rpt_family="Lx9"
repeat_region      6890..6921
                    /rpt_family="AT_rich"
repeat_region      8693..8735
                    /rpt_family="TTCC)n"
unsure             8699..8700
                    /note="probably CC"
repeat_region      9208..9234
                    /rpt_family="TTTG)n"
repeat_region      complement(9236..9335)
                    /rpt_family="PB1D7"
repeat_region      complement(9603..9751)
                    /rpt_family="B1F"
repeat_region      10571..10599
                    /rpt_family="AT_rich"
repeat_region      10637..10660
                    /rpt_family="AT_rich"
repeat_region      10951..11146
                    /rpt_family="B2_Mm1"
repeat_region      12671..12730
                    /rpt_family="TG)n"
repeat_region      13141..13214
                    /rpt_family="ID3"
repeat_region      13228..13256
                    /rpt_family="PB1D10"
repeat_region      13257..13323
                    /rpt_family="ID_B1"
repeat_region      13331..13358
                    /rpt_family="AT_rich"
repeat_region      13483..13667
                    /rpt_family="B3"
repeat_region      13769..14378
                    /rpt_family="L1M4"
repeat_region      14519..14619
                    /rpt_family="L1_MM"
repeat_region      complement(14772..15069)
                    /rpt_family="L1_MM"
repeat_region      15389..15420
                    /rpt_family="AT_rich"
repeat_region      17608..17670
                    /rpt_family="TATATG)n"
repeat_region      17773..17801
                    /rpt_family="AT_rich"
repeat_region      17825..17850
                    /rpt_family="AT_rich"
repeat_region      18110..18163
                    /rpt_family="AT_rich"
repeat_region      20210..20731
                    /rpt_family="L1_MM"
repeat_region      20759..20876
                    /rpt_family="RSINB1"
repeat_region      complement(21592..21799)
                    /rpt_family="B1_MM"
repeat_region      21818..21837
                    /rpt_family="(A)n"
repeat_region      21916..21949
                    /rpt_family="AT_rich"
repeat_region      complement(22545..22831)
                    /rpt_family="Lx7"
repeat_region      22892..23011
                    /rpt_family="B1_MM"
unsure             22949..22974
                    /note="single clone coverage"
```

```
repeat_region      complement(23095..23294)
                    /rpt_family="Lx7"
repeat_region      complement(23912..24049)
                    /rpt_family="B1_MM"
repeat_region      complement(24017..24236)
                    /rpt_family="MTC"
repeat_region      complement(24238..24355)

Alignment Scores:
Pred. No.:         1,11e-70
Score:             920.00
Percent Similarity: 98.94%
Best Local Similarity: 94.15%
Query Match:       95.63%
Db:                10
Gaps:              0

US-10-627-571-2 (1-188) x AC120859 (1-184327)

Qy      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeu 20
      :::
Db      144853 GTGGCTACAGATGCTTCATTCCAAAACCTGGCGCTTCAGGCACAAAAGAGATCCTG 144912

Qy      21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
Db      144913 GGCAAAATGCTATCCAAATCCATCCGCACACGCTGATCGACACACGACGCGAGGTG 144972

Qy      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLysGluAlaGluLys 60
      |||||
Db      144973 CTAGATGAGCTGTACAGGCTGACCAAGATGACCCAGAACAAAGAGGCGGAGAGG 145032

Qy      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAsnGln 80
      |||||
Db      145033 GTCATCAAGAACCTCATCAAGACGCTCATCAAGCTGGCGCTTCACACGAAACATGCG 145092

Qy      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysGlyValHisGlnLeuAla 100
Db      145093 TTCAATCAAGACGAGCTGGCGCTCATGAGAAATTCAGAAAGAGAGTGCACGCTTGCC 145152

Qy      101 MetThrValIleValSerPheHisGlnValAspTyThrPheAspAsnValLeuSerArg 120
      |||||
Db      145153 ATGACGGTGTGAGCTTCCACCGAGTACGATGACCTTCGACCGCAATGTGCTGTCCAG 145212

Qy      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||||
Db      145213 CTGCGTGAACGAGTGCAGAGAGCTCCTACACGAGATCATTCAGCGCACCTTACCGCCAAG 145272

Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
      |||||
Db      145273 TCTCACGCGACGGGTAAATATATGTCTTGACCATTTTTCAGATTGATTTTGGCTGCC 145332

Qy      161 LeuTyraAsnProPheGluAsnPhelysProHisleuGlnLysLeuCysAspGlyIleAsn 180
      |||||
Db      145333 TTGTCAATACCTCTTGGAAAGTTTAACCTCATTCAGAAACCTTGGACGCGCATCAC 145392

Qy      181 LysMetLeuAspGluGluAsnIle 188
      |||||
Db      145393 AAAATGTTAGATGAAGAGACACTA 145416

RESULT 21
AC095257      237561 bp      DNA      linear      HTG 09-NOV-2002
LOCUS      Rattus norvegicus clone CH230-10A12, WORKING DRAFT SEQUENCE.
DEFINITION      AC095257
ACCESSION      AC095257
VERSION      AC095257.5 GI:24818022
KEYWORDS      HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
      Rattus.
REFERENCE      1 (bases 1 to 237561)
      Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
      Allen,C., Allen,H., Albrooks,S., Amin,A., Angilano,D.,
```

Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismalo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, Y., Chen, Y., Chen, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Day-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Diya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Garbajosa, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guevara, W., Gunaratne, P., Haaland, M., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hoques, M., Hollins, B., Howell, S., Huliyil, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, S., Johnson, B., Johnson, R., Jolivet, A., Karpach, S., Kelly, S., Kelly, S., Khan, Z., King, L., Koyar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, D., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensunewa, L., Louisesed, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Ming, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwachukwu, O., Okunolu, G., Olarnpusoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Piopert, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., Puz, L., Puz, M., Quito, J., Rachin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, M., Saverly, G., Scherer, S., Scott, G., Shamsan, S., Shen, H., Shetty, J., Shvartbeyn, A., Sison, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Treloar, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.

Direct Submission

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 237561)

Worley, K. C.

Direct Submission

Submitted (16-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 237561)

Rat Genome Sequencing Consortium

Direct Submission

Submitted (09-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 9, 2002 this sequence version replaced gi:23264563. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature

table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GDFL
Center clone name: CH230-10A12
----- Summary Statistics
Assembly program: Phrap, version 0.990329
Consensus quality: 229749 bases at least Q40
Consensus quality: 231209 bases at least Q30
Consensus quality: 232003 bases at least Q20
Estimated insert size: 232575; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
1 237561: contig of 237561 bp in length.
Location/Qualifiers
1. 237561
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-10A12"
1. 1156
/note="wgs_end_extension"
clone_end:Sp6
6006. 6586
/note="clone boundary"
clone_end:Sp6
site:BCORI
end_sequence:BH03075"
76320. 79576
/note="wgs contig"
complement(232422..233289)
/note="clone boundary"
clone_end:T7
site:BCORI
end_sequence:BH03071"
234103. 235440
/note="wgs_end_extension"
clone_end:T7"
235793. 237561
/note="wgs_end_extension"
clone_end:T7"
ORIGIN
Alignment Scores:
Pred. No.: 2,17e-70 Length: 237561
Score: 918.00 Matches: 177
Percent Similarity: 98.40% Conservative: 8
Best Local Similarity: 94.15% Mismatches: 3
Query Match: 95.43% Indels: 0
DB: 2 Gaps: 0
US-10-627-571-2 (1-188) x AC095257 (1-237561)
QY 1 MetAlaThrAspValPheAsnSerLySAenLeuAlaGlnAlaGlnIlySerIleLeu 20
:::|||||
Db 95045 GGGCGTACAGATGTTTAAATTCACAAAACCTGGCGCTTCAGGCAAAAAGAGATCTGG 95104
QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerGluVal 40

```

```

Db      95105 GGTAAATGATCCAAATCCATCGCCACCACTCATGATGACACGACGCGAGTC 95164
Qy      41 LeuAaPGLuLeuTyraArgValThrArgGluTyrrThrGlnAsnIleYsGluAaGluYs 60
Db      95165 CTGGAGAGCTGTACAGGGGTGACCAAGAGTACACCAAGAAAGAGGAGGACGAAG 95224
Qy      61 LysIleYsAsnLeuIleYsThrValIleYsLeuAlaIleLeuTyraArgAsnGln 80
Db      95225 GTCATCAAGACCTCATCAAGACGTCATCAAGCTGGCGTTTCCACAGAACATCG 95284
Qy      81 PheAsnGlnAaPGLuLeuAlaLeuMetGluYsPheYsIleYsValIleGlnLeuAla 100
Db      95285 TTCACACCAAGACGAGCTGGCTGTATGAGAGAAATTCAAGAAAGGTGCACGCGTGGCC 95344
Qy      101 MetThrValIleSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
Db      95345 ATGACCGGTGTAGCTTTACCAAGGTAGATCACCTTTGACCGCAACGTCGTCCAG 95404
Qy      121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaYs 140
Db      95405 CTGCTCAAGACGAGTCCGAGAGCTCTGACGACATCATTCAGCGCCACCTTACCGCCAG 95464
Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
Db      95465 TCTCAGCGAGCGGTTAATACGCTTTGATCATTTTTCAGATTGTGATTTTGGCTGCC 95524
Qy      161 LeuTyraAsnProPheGluYsAsnPheYsProHisLeuGlnIleYsLeuCyAspGlyIleAsn 180
Db      95525 TTGTAAATATCCCTTTGGAAATTTAAACCTTACAGAACTTTGTGACGCGATCAAC 95584
Qy      181 LysMetLeuAspGluGluAsnIle 188
Db      95585 AAAATGTTGGATGAAGAAACATA 95608

RESULT 22
LOCUS      BD149395 816 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.
ACCESSION  BD149395
VERSION     BD149395.1 GI:27855153
KEYWORDS   JP 2002191363-A/4238.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 816)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and use thereof
Patent: JP 2002191363-A 4238 09-JUL-2002;
HELIX RESEARCH INSTITUTE

COMMENT
OS Homo sapiens (human)
PN JP 2002191363-A/4238
PD 09-JUL-2002
PF 28-JUL-2000 JP 200280990
PI TOSHIO OTA, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAZUO
PI SAITO.
PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU,
PI KEIICHI NAGAI, TETSUO OTSUKI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
PC 10, C12P21/02, C12Q1/68//C12P21/08, G06F1/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT source 1. 816
Location/Qualifiers
1. 816
/organism="Homo sapiens (human)".
FEATURES
source 1. 816
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

```

```

ORIGIN
Alignment Scores:
Pred. No.: 9,416-73 Length: 816
Score: 916.00 Matches: 184
Percent Similarity: 97.89% Conservative: 2
Best Local Similarity: 96.84% Mismatches: 2
Query Match: 95.22% Indels: 2
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x BD149395 (1-816)
Qy      1 MetAlaTrnAaPValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIleYsIleLeu 20
Db      101 GTGGCCACAGATCTTTAATTCCAAAAACCTGGCGCTTACGGCACAAAGAGATCTTG 160
Qy      21 GlyYsMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
Db      161 GGTAAATGCTGTCCAAATCCATCCGCCACCACTTAATAGACGACACAAAGTGTAGGTCG 220
Qy      41 LeuAaPGLuLeuTyraArgValThrArgGluTyrrThrGlnAsnIleYsGluAaGluYs 60
Db      221 CTGATGAGCTGTACAGGTGACAGGAGTACACCCAAACAAAGAGAGGACAGAGAG 280
Qy      61 LysIleYsAsnLeuIleYsThrValIleYsLeuAlaIleLeuTyraArgAsnGln 80
Db      281 ATCATCAAGAACCTCATCAAGACACTCATCAAGCTGGCCATTCTTATAGAAATATCAG 340
Qy      81 PheAsnGlnAaPGLuLeuAlaLeuMetGluYsPheYsIleYsValIleGlnLeuAla 100
Db      341 TTAAATCAAGATGACACTGATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCT 400
Qy      101 MetThrValIleSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
Db      401 ATGACCGGTGTAGTTCCATCAGGTGATTAATCTTTGACCGGAAATGTTATCCAG 460
Qy      121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaYs 140
Db      461 CTGTTAATATGATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCCAG 520
Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
Db      521 TCACATGAGCGGGTAAATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 580
Qy      161 LeuTyraAsnProPheGluYsAsnPheYsProHisLeuGlnIleYsLeuCyAspGlyIleAsn 180
Db      581 TTGTAAATATCCCTTTGGGAATTTAAACCCCACTTACAAACATAAGTATGATCAAC 640
Qy      181 LysMet--LeuAaPGLuGluAsnIle 188
Db      641 AAAATGTTGGATGAAGAAACATA 666

RESULT 23
LOCUS      AX869333 816 bp DNA linear PAT 17-DEC-2003
DEFINITION Sequence 4238 from Patent EP1074617.
ACCESSION  AX869333
VERSION     AX869333.1 GI:40024196
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 816)
Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,
Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
Primer for synthesizing full-length cDNA and their use
Patent: EP 1074617-A 4238 07-FEB-2001;
Research Association for Biotechnology (JRP)

COMMENT
OS Homo sapiens (human)
PN 1. 816
PD
PF
PI
PI
PI
PC
C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/
PC 10, C12P21/02, C12Q1/68//C12P21/08, G06F1/30, C12N15/00, C12N5/00 CC
Primer for synthesizing full-length cDNA and use thereof FH Key
FT source 1. 816
Location/Qualifiers
1. 816
/organism="Homo sapiens"
/mol_type="unassigned DNA"

```

ORIGIN /db_xref="taxon:9606"

Alignment Scores:
Pred. No.: 9,41e-73 Length: 816
Score: 916.00 Matches: 184
Percent Similarity: 97.89% Conservative: 2
Best Local Similarity: 96.84% Mismatches: 2
Query Match: 95.22% Indels: 2
DB: Gaps: 0

US-10-627-571-2 (1-188) x AX869333 (1-816)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAglInuLysIleLeu 20
DB 101 GTGGCCACAGATGCTTAAATTCACAAAACCTGGCCCTTACAGCACAAGAGATCTTG 160
QY 21 GlyMetValSerLysSerIleAlaThrThreuleleAspAspThrSerSerGluVal 40
DB 161 GGTAAGATGGATGTCACAAATCCATCCACCACTTAATAGACACACAGATGAGGTG 220
QY 41 LeuAspGluLeuValThrArgGluValThrGlnAsnLysGluAgluLys 60
DB 221 CTGGATGAGCTTACACAGTGAACAGGAGTACACCAAAACAAGAGAGGAGAGAG 280
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuValArgAsnGln 80
DB 281 ATCATCAAGAACCTCATCAAGACAGTCAAGCTGCTTTTATAGAAATATCAG 340
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValIleGlnLeuAla 100
DB 341 TTTAATCAAGATGAGTGAAGATGATGAGAAATTAAGAAAGTTCATCAGCTTGT 400
QY 101 MetThrValAlaSerPheHisGlnValAspValThrPheAspArgAsnValLeuSerArg 120
DB 401 ATGACCGTGTCACTTCCATCAGTGCATTAATACCTTACCGGAATGCTTATCCAG 460
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleGlnArgHisLeuThrAlaLys 140
DB 461 CTGTAATATCAATGACAGAGATGCTGCACCAATCATTTAGCGGCCACTCAGTCCAG 520
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 521 TCACATGAGACGGGTAAATATGTTGATCATTTTCAATGTTGAATTTTGGCTGCC 580
QY 161 LeuValAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyLeuAsn 180
DB 581 TTCTATATATCTTTGGGAATTTTAAACCCCATTAACAAACTTAAGTGAATGATATCAG 640
QY 181 LysMet-LeuAspGluGluAsnIle 188
DB 641 AAAATGCTTGGATGAAGACACTA 666

RESULT 24
AJ720906 2009 bp mRNA linear VRT 30-SEP-2004
LOCUS Gallus gallus mRNA for hypothetical protein, clone 29h8.
DEFINITION
AJ720906
ACCESSION
AJ720906.1 GI:53136471
VERSION
ORF1.
KEYWORDS
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE
AUTHORS Caldwell,R.B., Kierzek,A.M., Arakawa,H., Bezzubov,Y., Zaim,J.,
Friedler,P., Kuter,S., Blagodatelski,A., Kostovska,D., Koter,M.,
Plachy,J., Carninci,P., Hayashizaki,Y., and Buerstedde,J.M.
TITLE Full-length cDNAs from bursal lymphocytes to facilitate gene
function analysis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2009)

AUTHORS Caldwell,R.B.
TITLE Direct Submission
JOURNAL Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
1, D-85764 Neuburg, GERMANY
Location/Qualifiers

FEATURES

source

1..2009
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="29h8"
/cell_type="lymphocyte"
/tissue_type="bursa"
/clone_lib="rikeni"
/dev_stage="2 weeks old"
1..2009
/locus_tag="RCJMB04_29h8"
1..188
/locus_tag="RCJMB04_29h8"
189..755
/locus_tag="RCJMB04_29h8"
/note="ORF1"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAG32565.1"
/db_xref="GI:53136472"
/translation="MATDVFNKSLAIQAKILGKNSKSIATTLIDPTSDVDEL
VYTKETYNKKKEAEKIKNLIRKIVKALIVRNPNODEIALMEKPKKQHLAKT
VVSFHOVDTPRPRLPSKLINCRRELLHQLIQRLTKSHGRNNVFDHRSCEFLAA
LYNPFQYKLHLQKLCGVNRLDEBNT"
756..>2009
/locus_tag="RCJMB04_29h8"

3' UTR

ORIGIN

Alignment Scores:

Pred. No.: 3.88e-69 Length: 2009
Score: 880.00 Matches: 169
Percent Similarity: 96.28% Conservative: 12
Best Local Similarity: 89.89% Mismatches: 7
Query Match: 91.48% Indels: 0
DB: Gaps: 0

US-10-627-571-2 (1-188) x AJ720906 (1-2009)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAglInuLysIleLeu 20
DB 189 ATGGCCACGAGATGCTTCAATTCACAAAAGCTTGCCATTCAGGCCACAGAAAGATCTT 248
QY 21 GlyLysMetValSerLysSerIleAlaThrThreuleleAspAspThrSerSerGluVal 40
DB 249 GGGAAATGGATCCAAAGTCAATAGCACTACTTATATGATGATATCCAGCACTGATGT 308
QY 41 LeuAspGluLeuValThrArgGluValThrGlnAsnLysGluAgluLys 60
DB 309 TTGATGAGCTTCAACAGTGAACGAAGATACCAACCAAAACAAGAGAGAGAG 368
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuValArgAsnGln 80
DB 369 ATCATTAAGAACTCTTAAATATAGTCCCAATTTGGCAATTTCTTACAGAAACATCA 428
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValIleGlnLeuAla 100
DB 429 TTTAATCAAGATGAATTCACATGATGAGAAAGTTCAAGAAAGAAAGTTATCAGTGGC 488
QY 101 MetThrValAlaSerPheHisGlnValAspValThrPheAspArgAsnValLeuSerArg 120
DB 489 AAGACGCTGCTCACTTCCATCAGGAGTATTAATCTTTGACACGAAATTTCTGTGCAAA 548
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 549 CTGTAATGACGTGAGAGAGCTTCAATCAGATCATTCAGGCTCACTCACTGAGAAA 608

QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
|||||
Db 609 TGGCATTGACGCTGTCAACATGCTTTGATTCATCTTCGATTTGAAATTTTGGCTGCC 668
QY 161 LeuTyrAsnProPheGlyAsnPhelysProHisLeuGlnIlyLeuCysAspGlyIleAsn 180
|||||
Db 669 TTGTATAAATCCCTTTGGACCTTATAAATCTTCAATTCACAAATCTTTGTATGTGTCTCAAC 728
QY 181 LysMetLeuAspGluGluAsnIle 188
|||||
Db 729 AGAATCTAGATGAGGGAACATA 752
RESULT 25
AK097884 1766 bp mRNA linear PRI 30-JAN-2004
LOCUS AK097884
DEFINITION Homo sapiens CDNA FLJ40565 fls, clone THYM02004688, highly similar
to Homo sapiens MDC-1.3 isoform 2 mRNA.
ACCESSION AK097884
VERSION AK097884.1 GI:21757783
KEYWORDS oligo capping; fls (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS 1
Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Makamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Oobayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shibatoru, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, M., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yoshida, M., Houchi, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Nomiyama, H., Sato, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiroo, M., Omori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoch, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702035
2
Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,
Arita, M., Musashino, K., Yuki, H., Hara, H., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Ota, T., Makamatsu, A., Ishii, S., Yamamoto, J.,
Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K.,
Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H.,
Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A.,
Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuho, Y., Nagai, K.
and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 1766)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7

COMMENT
Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomic@shri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
source
1..1766
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYM02004688"
/cfeature_type="CDS"
/clone_1b="THYM02"
/note="Cloning vector: pME18FLJ3"
ORIGIN
Alignment Scores:
Pred. No.: 4.25e-67 Length: 1766
Score: 856.50 Matches: 173
Percent Similarity: 92.06% Conservative: 1
Best Local Similarity: 91.53% Mismatches: 1
Query Match: 89.03% Indels: 14
DB: 9 Gaps: 1
US-10-627-571-2 (1-188) x AK097884 (1-1766)
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeu 20
Db 176 ATGGCCACGATGCTTTTATTCCTCAAAACCTGCGCGTTCAGGCCACAAAGAGATCTTG 235
QY 21 GILYsMetValSerLysSerIleAlaThrIleLeuIleAspAspThrSerGluVal 40
Db 236 GGTAAATGTGTCCAAATCCATCCGACCACTTATATGACGACAAAGTAGTGAGGTG 295
QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLys 60
Db 296 CTGGATGAGCTTAAAGCA-----GAA 316
QY 61 Lys-IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
Db 317 GATCATCAAGAACCTTCATAGACAGTCTCAAGCTGGCCATCTTTATAGAAATATATCA 376
QY 80 nPheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAl 100
Db 377 GTTTATCAAGATGAGCTAGCTAGCTATGAGAGAAATTTAAGAAAGATTCAAGCTTGC 436
QY 100 aMetThrValValSerPheHisGlnValAspTyrThrPheAspAspValLeuSerAr 120
Db 437 TATGACCGGTGGTCAGTTTCCATCAGCTGATATATCTTTGACCGGAATGTGTATCCAG 496
QY 120 GLeuLeuAsnGluLysCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db 497 GCTGTAAATGATGAGAGAGATGTCGACCAAAATCTTACGCCGACCTCACTGCACAA 556
QY 140 sSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaI 160
Db 557 GTCACATGAGACGGGTAAATATGCTGTTCATCATTTTTCAGATGTGAAATTTTGGCTGC 616
QY 160 aLeuTyrAsnProPheGlyAsnPhelysProHisLeuGlnIlyLeuCysAspGlyIleAs 180
Db 617 CTGTGATATATCCCTTTTGGGAATTTTAAACCCACATTACAAAATATGATGATGATCA 676
QY 180 nLysMetLeuAspGluGluAsnIle 188
Db 677 CAAATGTTGATGAGAGAACATA 701
RESULT 26
CR760636

LOCUS CR760636 1559 bp mRNA linear VRT 15-SEP-2004
 DEFINITION Xenopus tropicalis finished cDNA, clone TEG9053008.
 ACCESSION CR760636
 VERSION CR760636.1 GI:51966545
 KEYWORDS Xenopus tropicalis (Silurana tropicalis)
 SOURCE Xenopus tropicalis
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus; Silurana.
 REFERENCE 1 (bases 1 to 1559)
 AUTHORS Amaya, E., Ashurst, J.L., Bonfield, J.K., Croning, M.D.R., Davies, R.M., Francis, M.D., Garrett, N., Gilchrist, M.J., Graham, D.V., McLaren, S.R., Papalopulu, N., Rogers, J., Smith, J.C., Taylor, R.G., Voigt, J. and Zorn, A.M.
 TITLE Direct Submission
 JOURNAL Submitted (15-SEP-2004) Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: trop@sanger.ac.uk
 COMMENT Sanger Xenopus tropicalis EST/cDNA project.
 This sequence is from a Xenopus Gene Collection (XGC) library, from a library constructed by Aaron M. Zorn. cDNA was prepared from RNA extracted from eggs. EcoRI-NotI cut cDNA was then ligated into pCS107 with EcoRI at the 5' end and NotI at the 3' end.
 Vector: pCS107; Site 1: EcoRI; Site 2: NotI
 Host: Escherichia coli XL1-blue.
 FEATURES
 Location/Qualifiers
 1..1559
 /organism="Xenopus tropicalis"
 /mol_type="mRNA"
 /db_xref="taxon:8364"
 /clone="TEG9053008"
 /clone_1ib="XGC-egg"
 /dev_stage="egg"
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,4e-66 Length: 1559
 Score: 844.50 Matches: 161
 Percent Similarity: 96.30% Conservative: 21
 Best Local Similarity: 85.19% Mismatches: 6
 Query Match: 87.79% Indels: 1
 DB: 5 Gaps: 1
 US-10-627-571-2 (1-188) x CR760636 (1-1559)
 QY 1 MetAlaThrAspValPheAsnSerLySAsnLeuAlValGlnAlaGlnLySLeu 20
 Db 155 ATGGCAACAGATGCTTTCAGTCCAAAATCTGGCTGACAAAGCCAAAGAAAGATTCTT 214
 QY 21 GlyLysMetVal---SerLySerIleAlaThrLeuIleAspAspThrSerSerGlu 39
 Db 215 GGGAAAGATGGCTCATCAAAAGTACATAGCAACATCCCTTATGACATCAAAAGCGGTGAA 274
 QY 40 ValLeuAspGluLeuThrArgValThrArgGluThrGlnAsnLySValGlnAlu 59
 Db 275 GTTCTAGATGAAGTGAACATTGACGAGGAGTACACCAGTAAGAAAGATCCGAA 334
 QY 60 LysLyS11leLysAsnLeuIleLysThrVal11leLysLeuAla11leLeuThrArgAsnAsn 79
 Db 335 AAAGTTATAAAGAACTCATCAAAACCGTATCAAGCTGCACTCCTGTACAGAAATAT 394
 QY 80 GlnPheAsnGlnAspGluLeuAlaLeuMetGlnLysPheLysValLysValHisGlnLeu 99
 Db 395 CAGTTTAATGAGAAAGATGACCTCATGAGAGATTCMAAGAAAGGCCATCAGCTG 454
 QY 100 AlaMetThrValValSerPheHisGlnValAspThrThrPheAspArgAsnValLleuSer 119
 Db 455 GCCATACCTGTGCTGAGCTTTTATCAGGTGGAATATACTTTGACGAAAGCGCTTTC 514
 QY 120 ArgLeuLeuAsnGluLysArgGluMetLeuHisGln11le11leGlnArgHisLeuThrAla 139
 Db 515 AATTGTTAATGATGAGCGGAGGAGTGTCTCATGAGTATACAGGCGCATCTCAGGGCA 574

QY 140 LysSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 159
 Db 575 AAGTCTCAGGAGGATTAATTAACGTGTTGACACTTTCAATATGTAATTTGGCT 634
 QY 160 AlaLeuThrAsnProPheGlyAsnPhelyProHisLeuGlnLysLeuCysAspGlyLe 179
 Db 635 GCTCTTACATCCCTTTGGACCTTTCAAAAGACACTTCAGAGACTCTGTGAAGAGT 694
 QY 180 AsnLyMetLeuAspGluGluValen11le 188
 Db 695 AACAAATGCTGGACGAGCAATATATT 721
 RESULT 27
 LOCUS BC072904
 DEFINITION Xenopus laevis MGC80354 protein, mRNA (cDNA clone MGC:80354 IMAGE:5074412), complete cds.
 ACCESSION BC072904
 VERSION BC072904.1 GI:49115085
 KEYWORDS MGC.
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; Xenopodinae; Xenopus; Xenopus.
 REFERENCE 1 (bases 1 to 1193)
 AUTHORS Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W. and Richardson, P.
 TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus Initiative
 JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
 PUBMED 12454917
 REFERENCE 2 (bases 1 to 1193)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, D., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uedlin, T.B., Toshiyuki, S., Carland, P., Prange, C., Raha, S.S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalius, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 PUBMED 12477932
 REFERENCE 3 (bases 1 to 1193)
 AUTHORS Klein, S. and Gerhardt, D.S.
 TITLE Direct Submission
 JOURNAL Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
 REMARK NIH-MGC Project
 COMMENT Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Igor David
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcsc.bc.ca
 Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth

Featherstone, Malachi Griffith, Obi Griffith, Ran Gun, Nancy Liao, Kim Macdonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prihpu, Parvaneh Saeedi, Jr Santos, Angelique Scherach, Utsula Skalka, Duane Smallus, Jeff Scott, Miranda Tsai, George Yang, Jaquale Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLN at: <http://image.llnl.gov>
Series: IRM Plate: 152 Row: 5 Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarly but not identically to protein.

FEATURES

Source

```

/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="WGC:80354 IMAGE:5074412"
/tissue_type="Ovary, Xenopus"
/clone_lib="NIDH XGC_Ov1"
/lab_host="NIDH"
/note="vector: pCMV-SPORT6"
1..1193
/gene="WGC80354"
/db_xref="locusID:443963"
153..722
CDS

```

ORIGIN

```

/product="MGCC80354 protein"
/protein_id="AAH72904.1"
/db_xref="GI:49115086"
/db_xref="locusid:443963"
/translation="MADIDNFKKALVVOAKKILGMAKSIATSLIDPTSGVLEDE
LYQATRETQKQKAEKTNLKYVIAKALVIRNQQPBEESIGKPKRVKQVLA
TVASFQVEYEVFADHAKKIDRELTQVDRHTAKSGRVANNVDFNSCEFLA
LVNPFQYKSHQRLNGVGNKMLKDDNI"

```

Alignment Scores:

Pred. No.:	5.03e66	Length:	1193
Score:	842.50	Matches:	160
Percent Similarity:	99.24%	Conservative:	20
Best Local Similarity:	84.66%	Mismatches:	8
Query Match:	87.58%	Indels:	1
DB:	5	Gaps:	1

US-10-627-571-2 (1-188) x BC072904 (1-1193)

[illegible]

QY	120	ArgLeuLeuAsnGluCysArgGluMetLeuGlnGlnIleIleGlnAlaGlnHisLeuThrAla	139
Db	513	AAATGTGTAACAGAGTCAGGAGAGTGCTTACACGAGCATACAGCCCATCTCACGGCA	572
QY	140	LysSerHisGluArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla	159
Db	573	AAGTGCACGGACGGTGTATATACGTGTTCACCCCTTTCAAATGTGAATTTTAACT	632
QY	160	AlaLeuTyrAsnProPheGluAsnPheIysProHisLeuGlnIlyLeuCysAspGlyIle	179
Db	633	GCTCTGTACAAATCCCTTTGGACCTTACAAAAAACAACCTTCAGACACTGTAAACGAGTC	692
QY	180	AsnIysMetLeuAspGluGluAsnIle	188
Db	693	AAACAAATGCTGATGAAGACAAATATT	719

RESULT 28

LOCUS	BC052765	3986 bp	mRNA	linear	VRT 12-OCT-2000
DEFINITION	Danio rerio tumor necrosis factor, alpha-induced protein 8, mRNA (cdna clone MGC:55331 IMAGE:2601021), complete cds.				

VERSION

KEYWORDS: MGC.

ORGANISM

ORGANISM *Danio rerio*

REFERENCE AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

AUTHORS

AUTHORS

TITLE

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
PUBMED
12477932

REFERENCE

AUTHORS Director MGC Project

JOURNAL

JOURNAL Submitted (23-MAY-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

COMMENT

COMMENT

Email: cgabs@remail.nih.gov
Tissue Procurement: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
cDNA Library Preparation: Dr. Sumio Sugano and Dr. Ko-ichi Kawakami
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org
Anup Madan, Jessica Pahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whitting

Series: IRAK Plate: 101 Row: c Column: 15
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 41056158.
 Location/Qualifiers

FEATURES

source

```

1. 3386
/organism="Danio rerio"
/mol_type="mRNA"
/strain="AB"
/db_xref="taxon:7955"
/clone="MGC:55331 IMAGE:2601021"
/rna_type="whole body, adult, (one male and one female,
including unfertilized eggs)"
/clone_1lb="Sugano Kawakami zebrafish DNA"
/lab_host="DH10B"
/notes="Vector: pME18S-FL3"
1. 3386
/gene="tnfaip8"
/notes="synonyms: MGC55331, zgc:55331"
/db_xref="LOCUSID:393303"
196. 756
/gene="tnfaip8"
/codon_start=1
/product="tumor necrosis factor, alpha-induced protein 8"
/protein_id="AAH52765.1"
/db_xref="GI:31127058"
/db_xref="LOCUSID:393303"
/translat="MDSFSTKMLAQOKKMSKMTKVANLFDITSEVLDELTYR
VTEKYYRNKREAKQIKMLKMYVKKGLVKNQFNNNEELATVRRKKVHTLMTAV
SFYIDFTFDRVMSNLLNDCRELLHQAIRHDLTASSHARINIVFNHFDCCPLATLY
GPSEYVRGHQIKCEGVNMLDEGNL"

```

ORIGIN

Alignment Scores:

Pred. No.: 1.37e-52 Length: 3986
 Score: 698.00 Matches: 125
 Percent Similarity: 86.49% Conservative: 35
 Best Local Similarity: 67.57% Mismatches: 25
 Query Match: 72.56% Indels: 0
 DB: 5 Gaps: 0

US-10-627-571-2 (1-188) x BC052765 (1-3986)

```

QY 4 AspValPheAnsSerlySaenleuAaValGlnAlaGlnlySylleuGlylMet 23
DB 199 GACTCGTTGACGACCAAGAGCTTATGACGACCGGAGGATGACGATG 258
QY 24 ValSerlySerlleaThrThrleuIleAspAspThrSerSerGluValleuApglu 43
DB 259 GCGACCAAGCAAGTGGCCAACTCTTATGACGACCGGAGGATGACGAG 318
QY 44 LeuTyArgValThrArgGluTyThrGlnAsnlySylleuGlnlySylleu 63
DB 319 CTGACGAGTGCACCAAGAGTACACGCGCACCGCAGAGGCGCCAGAGATCATCA 378
QY 64 AsnleuIlelySylleuValleuIleleuTyArgAsnAnginPheAsnGln 83
DB 379 AACCTCATCAAGATGGTGTCAAGTTGGGCTCTTACCGCAACGGTCAATCAAC 438
QY 84 AspGluLeuAlaIleuMetGluTyPheIlelySylleuValIleGlnleuIleMetThrVal 103
DB 439 GAGGAGTACCGTTAGCGAGCGGCTTTCGTAAGAGGTGATCGCTGGGATGACGCT 498
QY 104 ValSerPheIleGlnValAspTyThrPheAspArgAsnValleuSerArgleuLeuAsn 123
DB 499 GTTAGCTTCAACAGATCGACTTCACTTGAATCGACGCGTCAAGTATCTACTCAAT 558
QY 124 GluCyArgGluMetleuIleGlnIleleuIleGlnArgIleleuThrAlaIleSylleu 143
DB 559 GACTGCGGTAACTTTCGACACGCGCATCAATCGGATCAACGCGAATCTCAACGCC 618
QY 144 ArgValAsnAsnValPheAspPheIlePheSerAspCysGluPheleuAlaIleuTyArgAsn 163
DB 619 CGAATCAACATGCTTCAATCAATTTGCGCGAATTTGACTTCTCGGACGCTAATCGGA 678

```

```

QY 164 ProPheGlyAsnPheIleProIleGlnlySylleuCyArgAspIleleuIleMetleu 183
DB 679 CCTTCAGAAAGTACCGCGGCACTTGACAAAGATCTGTGAAGAGTCAACAAATGCTG 738
QY 184 AspGluGluAsnIle 188
DB 739 GACGAAAGCATCTT 753
RESULT 29
BX649252 205949 bp DNA linear VRT 06-FEB-2004
BX649252/c
LOCUS
DEFINITION
Zebrafish DNA sequence from clone DKEY-49M19 in linkage group 22,
complete sequence.
ACCESSION
BX649252
VERSION
BX649252.7 GI:42454788
KEYWORDS
HTG.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 205949)
REFERENCE
AUTHORS
Martin,S.
TITLE
Direct Submision
JOURNAL
Submitted (06-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Feb 6, 2004 this sequence version replaced gi:41392451.
COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhifeng Bao and Sean Boddy, submitted), and those beginning 'dir' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see

http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-49M19 is from a zebrafish BAC library

VECTOR: pindigBAC-5.1 library

FEATURES

source

```

1. 205949
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_1lb="Dantokey"

```

ORIGIN

Db 564 CGCATCAACCAAGCTTTAAACAATTGACAGACTGATGTTCTTCGCGCCCTCTACGG 623

Qy 164 ProPheGlyAsnPhelysProHisLeuGlnIleuCyAspGlyIleAsnIleuMetLeu 183

Db 624 CCGAGCCGAGCCCTTACCGCACCTGGAAGAGATGTCGAGAGTGTGAACAAGATGCTG 683

Qy 184 AspGluGluAsnIle 188

Db 684 GAGAGAGACACACATG 698

RESULT 31

LOCUS CQ736656 1248 bp DNA linear PAT 03-FEB-2004

DEFINITION Sequence 22590 from Patent WO02068579.

ACCESSION CQ736656

VERSION CQ736656.1 GI:42332642

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W. Kites, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof WO 02068579-A 22590 06-SEP-2002;

JOURNAL PE Corporation (NY) (US)

FEATURES

source 1.1248

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 6,44e-49 Length: 1248

Score: 651.00 Matches: 114

Percent Similarity: 89.07% Conservative: 49

Best Local Similarity: 62.30% Mismatches: 20

Query Match: 67.67% Indels: 0

Gaps: 0

US-10-627-571-2 (1-188) x CQ736656 (1-1248)

Qy 4 AspValPheAsnSerIleAsnLeuAlaValGlnAlaGlnIleuIleuGlyIleuGlyMet 23

Db 691 GATGTTTAACTTCAAGAGCTCTTGCGCTTCAAGCCAGAGAAAGATTCTGACCAAAATA 750

Qy 24 ValSerIleSerIleAlaThrThrIleuIleAspAspThrSerSerGluValIleuAspGlu 43

Db 751 GCCAGCAAAACTGTGCCCAACATGTTGATGATGACACCGACGACGATCTTGTATGAG 810

Qy 44 LeuTyArgValThrArgIleuTyThrGlnAsnIleuGlnAlaGlnIleuIleuGlyIleuGly 63

Db 811 CTCTCAAAAGTCCCAAGAGCACACACACAAAGAGAAAGCCCAAGATCATGAAA 870

Qy 64 AsnIleuIleuTyThrValIleuIleuAlaIleuTyArgAsnAsnGlnPheAsnGln 83

Db 871 GACTTAATCAAGGTGCGATCAAAATCGGGATCTCTACCGGAACACGATTACCCAA 930

Qy 84 AspGluLeuAlaIleuMetGluTyPheIleuValIleuGlnIleuAlaMetThrVal 103

Db 931 GAGGAGCTGTTATTTGTGGAAGATTCGGAGAAAGCTGAACACGACCGCATGACCAT 990

Qy 104 ValSerPheHisGlnValAspTyThrPheAspArgAsnValIleuSerArgIleuAsn 123

Db 991 GTCAAGCTTCTATAGAGTGAATACACCTTCGATGAGAACCTGCTCTCCAAATCTCTGCAT 1050

Qy 124 GluCyArgGluMetLeuHisGlnIleuIleuGlnIleuArgHisLeuThrAlaIleuSerHisGly 143

Db 1051 GAGTCAAGAGACCTGTGATGAACGTGTGACGCGGACCTGACGCGCCAGACCCAGCGG 1110

Qy 144 ArgValAsnAsnValPheAspHisPheSerAspCyAspGluPheLeuAlaIleuTyArg 163

Db 1111 CGCATCAACCAAGCTTTAAACAATTGACAGACTGATGTTCTTCGCGCCCTCTACGG 1170

Qy 164 ProPheGlyAsnPhelysProHisLeuGlnIleuCyAspGlyIleAsnIleuMetLeu 183

Db 1171 CTGATGAGAGACTGTGAGCCCAACCTCAAGAGATTTGTGAAGATCATTAAGTTGCTA 1230

Qy 184 AspGluGlu 186

Db 1231 GATGAGAAA 1239

RESULT 32

LOCUS CQ841622 1986 bp DNA linear PAT 02-AUG-2004

DEFINITION Sequence 269 from Patent EP1440981.

ACCESSION CQ841622

VERSION CQ841622.1 GI:50893409

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Iwagai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Nagai, K. and Irie, R. Full-length human cdna Patent: EP 1440981-A 269 28-JUL-2004;

JOURNAL Research Association for Biotechnology (JP)

FEATURES

source 1.1986

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 1,04e-48 Length: 1986

Score: 651.00 Matches: 114

Percent Similarity: 89.07% Conservative: 49

Best Local Similarity: 62.30% Mismatches: 20

Query Match: 67.67% Indels: 0

Gaps: 0

US-10-627-571-2 (1-188) x CQ841622 (1-1986)

Qy 4 AspValPheAsnSerIleAsnLeuAlaValGlnAlaGlnIleuIleuGlyIleuGlyMet 23

Db 152 GATGTTTAACTTCAAGAGCTCTTGCGCTTCAAGCCAGAGAAAGATTCTGACCAAAATA 211

Qy 24 ValSerIleSerIleAlaThrThrIleuIleAspAspThrSerSerGluValIleuAspGlu 43

Db 212 GCCAGCAAAACTGTGCCCAACATGTTGATGATGACACACGACGACGATCTTGTATGAG 271

Qy 44 LeuTyArgValThrArgIleuTyThrGlnAsnIleuGlnAlaGlnIleuIleuGlyIleuGly 63

Db 272 CTCTCAAAAGTCCCAAGAGCACACACACAAAGAGAAAGCCCAAGATCATGAAA 331

Qy 64 AsnIleuIleuTyThrValIleuIleuAlaIleuTyArgAsnAsnGlnPheAsnGln 83

Db 332 GACTTAATCAAGGTGCGATCAAAATCGGGATCTCTACCGGAACACGATTACCCAA 391

Qy 84 AspGluLeuAlaIleuMetGluTyPheIleuValIleuGlnIleuAlaMetThrVal 103

Db 392 GAGGAGCTGTTATTTGTGGAAGATTCGGAGAAAGCTGAACACGCGCATGACCAT 451

Qy 104 ValSerPheHisGlnValAspTyThrPheAspArgAsnValIleuSerArgIleuAsn 123

Db 452 GTCAAGCTTCTATAGAGTGAATACACCTTCGATGAGAACCTGCTCTCCAAATCTCTGCAT 511

Qy 124 GluCyArgGluMetLeuHisGlnIleuIleuGlnIleuArgHisLeuThrAlaIleuSerHisGly 143

Db 512 GAGTCAAGAGACCTGTGATGAACGTGTGACGCGGACCTGACGCGCCAGACCCAGCGG 571

QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeuTyrAsn 163
 572 CGCATCAACACGCTCTTTAACACCTTGGCGATGCGAGTCTCTCCACCCCTCTTAAT 631
 Db
 QY 164 PropheGlyAsnPhelyProHisLeuGlnLysLeuCyaspGlyLeuAsnLysMetLeu 183
 632 CTGCGATGAGACTGTAGCCCAACCTCAAGAGATTGTGAAGATCAATAGTTGCTA 691
 Db
 QY 184 AspGluGlu 186
 692 GATGAGAA 700
 Db
 RESULT 33
 AK123281 1986 bp mRNA linear prt 09-SEP-2003
 LOCUS Homo sapiens cDNA FL41287 f15, clone BRAMY2038484, moderately
 DEFINITION similar to Homo sapiens MDC-3.13 isoform 2 mRNA.
 ACCESSION AK123281
 VERSION AK123281.1 GI:34528784
 KEYWORDS oligo capping; f15 (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1
 AUTHORS Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
 Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
 Yamamoto, U., Isono, Y., Kawai, Hio, Y., Saito, K., Nishikawa, T.,
 Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
 Kikuchi, H., Kanda, K., Magatsuma, M., Murakawa, K., Kanehori, K.,
 Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y.,
 Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 TITLE Unpublished
 JOURNAL 2 (bases 1 to 1986)
 REFERENCE Isogai, T. and Yamamoto, U.
 AUTHORS Direct Submission
 JOURNAL Submitted (15-JUL-2003) Takao Isogai, Flu Project (HRI Team); 2-6-7
 Kazusa-Kametari, Kisarazu, Chiba 292-0818, Japan
 COMMENT (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction: Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.
 FEATURES
 source
 1..1986
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="BRAMY2038484"
 /cissue_type="amygdala"
 /clone_lib="BRAMY2"
 /note="Cloning vector: pME18SFL13"
 95..709
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="BAC85572.1"
 /db_xref="GI:34528785"
 /translation="MDSDSGSGSEGPYTAAGDPVSSKSLAQAOKTILSTASTV
 ANMLIDTSSSEIPDELKYTKKTHAKKAAHKIMKIDLYATIKGILYRNKPSQSEL
 VIVEKRRKLNGTAMTIVSFYEVEYFDNANVSNLLHECKDLVHELIVQSHLPTRTGR
 INHVFHFADVEFLSTLYSLDDDCRNLRICEGINKLDEKVL"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,04e-48 Length: 1986
 Score: 651.00 Matches: 114
 Percent Similarity: 89.07% Conservative: 49

Best Local Similarity: 62.30% Mismatches: 20
 Query Match: 67.67% Indels: 0
 DB: 9 Gaps: 0
 US-10-627-571-2 (1-1986) x AK123281 (1-1986)
 QY 4 AspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysMet 23
 152 GATGTTTATTAGTTCAAGAGCTCTTCCGCTTCAAGCCAGAGAGATTCGAGCAAAATA 211
 Db
 QY 24 ValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGlu 43
 212 GCCAGCAAAACCTGTGGCCCAATGTTGATGTGACACACGACGAGATCTTTATATAG 271
 Db
 QY 44 LeuTyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLysLysIleLys 63
 272 CTCTCAAGAGTACCAAGAGACACACACACACAGAGAGAGCCACAGATCTAGTAA 331
 Db
 QY 64 AsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGln 83
 332 GACTTAATCAAGTGGCGCATCAAAATCGGATCTCTACCGGACACACAGTTTAGCCA 391
 Db
 QY 84 AspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThrVal 103
 392 GAGGAGCTGTATTATGTGAGAAAGTTCCGAAAGAGCTGAAACAGACCCCATGACATT 451
 Db
 QY 104 ValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsn 123
 452 GTCAAGCTTCTATAGAGTGGATGCAATACCTTCGATAGAAAGTCTCTCCATCTCTGCAT 511
 Db
 QY 124 GlyCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGly 143
 512 GAGTGCAGAGACCTGGTGATGAATCGATGAGCGGACCTGACCCCGACGACCCAGCGG 571
 Db
 QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeuTyrAsn 163
 572 CGCATCAACACGCTCTTTAACACCTTGGCGATGCGAGTCTCTCCACCCCTCTTAAT 631
 Db
 QY 164 PropheGlyAsnPhelyProHisLeuGlnLysLeuCyaspGlyLeuAsnLysMetLeu 183
 632 CTGCGATGAGACTGTAGCCCAACCTCAAGAGATTGTGAAGATCAATAGTTGCTA 691
 Db
 QY 184 AspGluGlu 186
 692 GATGAGAA 700
 Db
 RESULT 34
 AC012678 188937 bp DNA linear HTG 18-JUL-2000
 LOCUS Homo sapiens chromosome 15 clone RP11-394B5, WORKING DRAFT
 DEFINITION SEQUENCE, 10 unordered pieces.
 ACCESSION AC012678
 VERSION AC012678.2 GI:6728953
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 188937)
 AUTHORS Gluhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
 Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelm, J.,
 Yu, S. and Davis, R.W.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 188937)
 AUTHORS Bruno, D., Conn, L., Dela Rosa, M., Federspiel, N., Foreman, P.,
 Gluhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
 Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Wilhelm, J.,
 Yu, S. and Davis, R.W.
 TITLE Direct Submission
 JOURNAL Submitted (03-NOV-1999) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA

Center code: UMWSC
Web site: http://chroma.mbt.washington.edu/msg_www
Contact: leerowen@systemsbiology.org
Drafting center: SDSTDC

----- Summary Statistics -----
Sequencing vector: pUC18; L08752
Chemistry: Dye-terminator Big Dye; 90% of reads
Chemistry: Dye-terminator Big Dye; 10% of reads
Assembly program: Phrap; version 0.990399

Note: Date from overlapping BACs AC022407 (drafting center: UMWSC) and AC012169 (drafting center: UMWSC) were added for finishing

FEATURES

source

1. 189796
Location/Qualifiers

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15q21.2"

/clone_id="RP11-394B5"

/note="Data from overlapping BACs RP11-105D1 and RP11-522G20 were added and the consensus sequence was determined from RP11-394B5 to the extent possible"

misc_feature

1. 15788
/note="Overlap with RP11-105D1 AC022407"

unsure

17935..17940
/note="low quality data"

misc_feature

55786..56417
/note="sequence data generated from subcloned PCR product only"

unsure

55786..56417
/note="low quality data"

unsure

65115..65120
/note="low quality data"

unsure

104690..104715
/note="low quality data"

misc_feature

105973..189796
/note="Overlap with RP11-522G20 AC012169"

unsure

161495..161500
/note="low quality data"

ORIGIN

Alignment Scores:

Pred. No.: 1,158-46

Score: 651.00

Percent Similarity: 89.07%

Best Local Similarity: 62.30%

Query Match: 67.67%

DB: 9

US-10-627-571-2 (1-188) x AC073964 (1-189796)

QY 4

AspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMet

DB 68379

GATGTTTAACTTCAAGAGAGCTTCCGCTCAAGCCAGAAAGAAATTCAGCAAAATA

QY 24

ValSerLysSerLysAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGlu

DB 68319

GCCACAAACTGTGGCCAACTGTGTGATGTGACACACGACGACGAGATCTTTATAG

QY 44

LeuTYrAspValThrArgGluValThrGlnAsnLysGluAlaGluLysLysIleLys

DB 68229

CTTTCACAAAGTCACAAAGACACACACAAAGAAAGCCACAAAGATCTTGACCA

QY 64

AsnLeuIleLysThrValIleLysLeuAlaIleLeuTYrArgAspAsnGlnPheAsnGln

DB 68139

GACTTAATCAAGCGTGGCATAAATCGGATCTCTACCGGACAAACCAAGTTTGCCA

QY 84

AspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAlaMetThrVal

DB 68139

GAGGAGCTGTATTGTGGAGAAATTCCGGAAGAGCTGAACCAAGCCGCTGACCAT

QY 103

68139

QY 104 ValSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArgLeuLeuAsn 123

DB 68079 GTCACTTTCTATGAGCGTGGATACACTTGCATATGGAACGCTCTCCAACTCTGCAT 68020

QY 124 GluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGly 143

DB 68019 GAGTGCAAGACCTGTGTGATGAACTGATGAGGAGGACCTGACGCCACCAAGCAACGCG 67960

QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysArgGluPheLeuAlaLeuTYrAsn 163

DB 67959 GCATCAACCACTGCTTAACTTGCATGATGAGGATGAGTCTCTCCACCTCTAATAGT 67900

QY 164 ProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeu 183

DB 67899 CTGATGAGACCTGTAGGCCACCTCAAGAGATTTGTGAAGAAATCAATATGCTTA 67840

QY 184 AspGluGlu 186

DB 67839 GATGAGAA 67831

RESULT 36

AC115187_1/c

WPCOMMENT

Sequence split into 5 fragments LOCUS AC115187 Accession AC115187

Fragment Name Begin End

AC115187_1 1 110000

AC115187_2 200001 310000

AC115187_3 300001 410000

AC115187_4 400001 420106

Continuation (2 of 5) of AC115187 from base 100001 (AC115187 Rattus norvegicus clone CH2

US-10-627-571-2 (1-188) x AC115187_1 (1-110000)

QY 5

ValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMetVal 24

DB 32867 GTGTTTAACTTCAAGAGAGCTTCCGCTCAAGCCAGAAAGAAATTCAGCAAAATA 32808

QY 25

SerLysSerLysAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeu 44

DB 32807 AGCAAAACTGTGGCCAACTGTGTGATGTGACACACGACGACGAGATCTTTATAG 32748

QY 45

TyrArgValThrArgGluValThrGlnAsnLysGluAlaGluLysLysIleLysAsn 64

DB 32747 TACAAAGTCACCAAGAAACCAACCAAGAAAGAGGCCCAAGATCATAGAGAT 32688

QY 65

LeuIleLysThrValIleLysLeuAlaIleLeuTYrArgAsnGlnPheAsnGlnAsp 84

DB 32687 GTAATTAAGGTGGCGATCAAAATGCTTCTTACCGGACAAACAGTCTTATAGTCAGAG 32628

QY 85

GluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAlaMetThrValVal 104

DB 32627 GAGGTGTGATGTGAGAAATCTCGGAAAGAAATGAAACCAAGCCGATATGATGATGATG 32568

QY 105

SerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGln 124

DB 32567 AGCTTCTACGAAGTGGAGTAACTTGTATCGAATGCTCTTAAACTTCTCATGAG 32508

QY 125

CysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArg 144

DB 32507 TGCAGGACCTGTGCTCATGAACTGTACAGGACACTGAGGCCCAAGAACCAAGGAGGAGC 32448

QY 145

ValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTYrAsnPro 164


```

DB      32447 ATAAACCATGTCTTCAACACCTTCGCTGATGATGATTCCTTTCCACTCTCAAGCTCGG 32388
QY      165 PheGlyAanPheIlyProHiIleuGlnIlyIleuGAspGlyIleuAnIlySHeIleuAap 184
DB      32387 CACGGAACTGACGCCCATCTCAAGAGATTTCGAAGATCAATCAATTAATTTGATGAT 32328
QY      185 GIUGlu 186
DB      32327 GACAAA 32322

RESULT 37
AC115505/c
LOCUS   Rattus norvegicus clone CH230-77C4, *** SEQUENCING IN PROGRESS ***,
DEFINITION
AC115505
AC115505
AC115505.4 GI:25188467
HTG: HTGS PHASE1: HTGS_DRAFT: HTGS_ENRICHED.
KEYWORDS
SOURCE  Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 256285)
Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsdorfs, S., Amin, A., Anguiano, D.,
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Bivental, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
DeGado, O., Denison, S., Deramo, C., Ding, Y., Din, H., Divya, K.,
Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,
Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,
Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K.,
Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, V.,
Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, W.,
Hollins, B., Howells, S., Huliy, S., Hume, J., Idelbitz, D., Jackson, A.,
Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolyvet, A.,
Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
Kowitz, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,
Lorenshuwa, L., Louisedge, H., Lozado, R. J., Lu, X., Ma, J.,
Madeshwar, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,
Mawliny, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,
Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nat, L.,
Nankervis, C., Neal, D., Newton, S., Nguyen, N., Norris, S.,
Nwokolelemeh, O., Okwomou, G., Olarnpunaagoon, A., Pal, S., Parks, K.,
Pascherak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,
Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L. L.,
Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R.,
Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smaj, D.,
Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Soza, J.,
Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
Valdes, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczyk, R., Woodson, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhou, S., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Unpublished
2 (bases 1 to 256285)
Worley, K. C.
Submitted (20-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 256285)
Rat Genome Sequencing Consortium.
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 23, 2002 this sequence version replaced gi:23269911.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSWY
Center clone name: CH230-77C4
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 199255 bases at least Q40
Consensus quality: 203424 bases at least Q30
Consensus quality: 206379 bases at least Q20
Estimated insert size: 195822; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
24557: contig of 24557 bp in length
24558
24657: gap of unknown length
24658
201810: contig of 17715 bp in length
201811
201910: gap of unknown length
201911
214313: contig of 12403 bp in length
214314
214413: gap of unknown length
214414
247000: contig of 32587 bp in length
247001
247101: gap of unknown length
247102
248513: contig of 1413 bp in length
248514
248613: gap of unknown length
248614
249865: contig of 1373 bp in length
249867
250086: gap of unknown length
250087
251489: contig of 1403 bp in length
251490
251589: gap of unknown length
251590
252936: contig of 1347 bp in length
252937
253036: gap of unknown length
253037
256285: contig of 3249 bp in length.
Location/Qualifiers
1..256285
/organism="Rattus norvegicus"

```

```

/mol type="genomic DNA"
/db xref="taxon:10116"
/clone="CH230-77C4"
misc_feature
1..1142
  /note="wgs_contig"
misc_feature
4086..6215
  /note="wgs_contig"
misc_feature
24658..25854
  /note="wgs_contig"
misc_feature
228541..229886
  /note="wgs_contig"
misc_feature
245270..247000
  /note="wgs_contig"
ORIGIN
Alignment Scores:
Pred. No.: 6.3e-45 Length: 256285
Score: 633.00 Matches: 109
Percent Similarity: 87.91% Conservatave: 51
Best Local Similarity: 59.89% Mismatches: 22
Query Match: 65.80% Indels: 0
DB: 2 Gaps: 0
US-10-627-571-2 (1-188) x AC115505 (1-256285)
Oy 5 ValPheAenSerLySAnLeuAlaValGlnAlaGlnLySAlLeuGlyLeuMetVal 24
Db 120592 GTGTTACTCCAGAGAGTCCGCTTCAAGCCCGAAGAGATCTTGAAGCAAGTACG 120593
Oy 25 SerLySeriLeAlaThrThrLeuLeaApeThrSerSerGlyValLeuApeGluLeu 44
Db 120532 AGCAAAACCTGGGCGCAACATGTCATGATGACACCAAGAGAGATCTTGCATGAGCTG 120473
Oy 45 TyrArgValITrrArgGluTyrThrGlnAnuLySylGluAlaGluLySylLeysAn 64
Db 120472 TACAAAGTCACCGAAGAACACACCCCAAGAGAGGCCACAGATCATGAAAGAT 120413
Oy 65 LeuileYsThrValilLeuYlLeuYlLeuYlYrAgsAnSnglNpheAnSnglAn 84
Db 120412 GTATTAAGAGTGGGCATCAAAATTGTTATCTCTACCGAAGAACAGTTCATGAAG 120353
Oy 85 GluLeuAlaLeuMetGluLySylLeuYlLeuYlValHsGlnLeuAlaMetThrValVal 104
Db 120352 GAGGTTCAGTATGGTGGAGAAACTCCGAGAAAGAAACGAGACCGCATGACATGTC 120293
Oy 105 SerPheHsGlnValAspTyrThrPheApeAgsAnValLeuSerGlyLeuLeuAnSglu 124
Db 120292 AGCTTCTACGAGAGTGAAGTATACCTTTGATACGAATGCTCTTAAACTTCATGATGAG 120233
Oy 125 CysAArgIuMetLeuHsGlnIleleGlnAtrGhsleuThrAlaLySerHsGlyArg 144
Db 120232 TCGCAAGACCTGGTCCATGAAGACTGTTACAGCACACTTGACGCCGAGAACCCAGGAGCC 120173
Oy 145 ValAenAanValPheApeHsApeSerApeCysGluPheLeuAlaAlaLeuYrAanPro 164
Db 120172 ATAAACCATGTCCTCAACCATCTCGTATGGAATTCCTTCAACCTCTACAGTCCG 120113
Oy 165 PheGlyAenPheLySylPheHsleuGlnLySylCysApeGlylLeuAnLyMetLeuApe 184
Db 120112 CACGGAACCTGACAGCCCAATCTCAAGAGGATTTGCGAAGGAAATCAATAATTGTTAAT 120053
Oy 185 GluGlu 186
Db 120052 GACCAA 120047
RESULT 38
AC129440 273225 bp DNA linear HTG 13-MAY-2003
LOCUS Rattus norvegicus clone CH230-112K14, *** SEQUENCING IN PROGRESS
DEFINITION *** 7 unordered pieces.
ACCESSION AC129440
VERSION AC129440.3 GI:30581254
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

```

SOURCE	ORGANISM
Rattus norvegicus (Norway rat)	
Rattus norvegicus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;	
Rattus.	
1 (bases 1 to 273225)	
Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,	
Allen,C., Allen,H., Alebrooks,S., Amin,A., Angiano,D.,	
Ayalebechi,V., Aoyagi,A., Ayodeji,M., Bacc,E., Baden,H.,	
Baldwin,D., Bandaranaike,D., Barber,M., Barnesread,M., Benahmed,F.,	
Biswal,O,K., Blair,D., Blankenburg,K., Blyth,P., Brown,M.,	
Bryant,N., Buhay,C., Burch,P., Butrell,K., Calderon,E.,	
Cardenas,V., Carter,K., Cavazos,I., Caesar,H., Cente,A.,	
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,	
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,	
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,	
Deigado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Diya,K.,	
Draper,H., Dugan-Rocha,S., Dunn,A., Dudin,K., Duval,B., Eaves,K.,	
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,	
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,	
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,	
Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,	
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,	
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,	
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,	
Hollins,B., Howells,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A.,	
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,	
Kapathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Koyat,C.,	
Kowis,C., Kraft,C.L., Lebow,H., Levan,D., Lewis,L., Li,Z., Liu,J.,	
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,	
Lorenshewa,L., Louissege,H., Lozado,R.J., Lu,X., Ma,J.,	
Maheshwari,M., Mahindartne,M., Mahmud,M., Malloy,K., Mangum,A.,	
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,	
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,	
Miliosavilevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,	
Morgan,M., Morris,K., Morris,S., Muidessa,M., Murphy,M., Natr,L.,	
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,	
Nwokediameh,O., Okwoum,G., Olarunsgoon,A., Pal,S., Parke,K.,	
Paaternak,S., Paul,H., Perez,A., Perez,L., Pfannoch,C.,	
Plopper,P., Poindexter,A., Popovic,D., Primus,E., Pu,L.,L.,	
Puazo,M., Quiroz,J., Rachlin,B., Reeves,K., Regier,M.A., Relgh,R.,	
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,	
Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,	
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,	
Shetty,J., Shvartbeyn,A., Sisson,I., Sitter,C.D., Snajs,D.,	
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,	
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabot,P., Taylor,C.,	
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,	
Valas,R., Vera,Y., Villasaana,D., Waldron,L., Walker,B., Wang,J.,	
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,	
Williams,G., Willson,R., Wleczky,R., Wooden,H., Worley,K.,	
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,	
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von	
Niederhausem,A., Weise,R., Smith,D.R., Holt,R.A., Smith,H.O.,	
Weinstock,G. and Gibbs,R.A.	
Direct Submission	
Unpublished	
2 (bases 1 to 273225)	
Worley,K.C.	
Direct Submission	
Submitted (30-JUL-2002) Human Genome Sequencing Center, Department	
of Molecular and Human Genetics, Baylor College of Medicine, One	
Baylor Plaza, Houston, TX 77030, USA	
3 (bases 1 to 273225)	
Rat Genome Sequencing Consortium.	
Direct Submission	
Submitted (13-MAY-2003) Human Genome Sequencing Center, Department	
of Molecular and Human Genetics, Baylor College of Medicine, One	
Baylor Plaza, Houston, TX 77030, USA	
On May 13, 2003 this sequence version replaced gi:22364293.	
The sequence in this assembly is a combination of BAC based reads	
and whole genome shotgun sequencing reads assembled using Atlas	
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described	

in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GKTU

Summary Statistics

Assembly program: Atlas 3.0!
Consensus quality: 224959 bases at least Q40
Consensus quality: 227786 bases at least Q30
Consensus quality: 229294 bases at least Q20
Estimated insert size: 236137: sum-of-contigs estimation
Quality coverage: 6x in Q20 bases: sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES

*	1	4842:	conf: of 4842 bp in length
*	4843	4992:	gap of unknown length
*	4943	265126:	conf: of 260184 bp in length
*	265127	265226:	gap of unknown length
*	265227	266729:	conf: of 1503 bp in length
*	266730	266829:	gap of unknown length
*	266830	266869:	conf: of 2040 bp in length
*	268970	268969:	gap of unknown length
*	268970	270137:	conf: of 1158 bp in length
*	270128	270227:	gap of unknown length
*	270228	271518:	conf: of 1291 bp in length
*	271519	271618:	gap of unknown length
*	271619	273225:	conf: of 1607 bp in length.
		Location/Qualifiers	
		1..273225	

1.27322

```
misc_feature
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-112K14"
    complement(1198..4940)
    /note="clones_boundary"
    clone_end:17
    site_50cpi
    end_sequence:BH258362"
    4943..6894
    /note="wgs_config"
misc_feature
```

ORIGIN

Alignment Scores:	
Pred. No.:	6,73e-45
Score:	633.00
Percent Similarity:	87.91%
Best Local Similarity:	59.89%
Query Match:	65.80%
DB:	2
	Caps: 0
	Gaps: 0
	Length: 273222
	Matches: 109
	Conservative: 51
	Mismatches: 22
	Indels: 51

US-10-627-571-2 (1-188) x AC129440 (1-273225)

OY		5	ValphaanserlysaamluA1aValGlnabaglnlslslleluGlyMeVai	24
Db	8669	GTTTATAGCTCCAAAGACTTCGCCCTTCAACCGCAAGAAGATCTTGACAATAATGCC	8728	
OY		25	SerlysserilIealathrThrLeuileAspaapThrsSerSergluValleAspGluDeu	44
Db	8729	AGCAAACCTGGGCACACGTCTCATGATGTATACACGACCGAGATCTTTGATGACGTG	8788	
OY		45	TyrarValIthrarglutyrrThGlnasnlslslglualaglulslsllelyasn	64
Db	8789	TACAAAGTCACCGAAGAACACACCACAACAAAGAGAGGCCCAAGATCATGAAGAT	8848	
OY		65	leuilelsythrValllelyslsuAlilleuulyrrgrasnaenGlnphasnglasp	84
Db	8849	GTAATTAAGGTGGCGATCAAAAATTGGTATTTCTTACCGGAACAAACAGTTTAGTCAGAG	8908	
OY		85	GluUeuAlaleuWeecGluLyPheLysLysValhisGlnleuAlamethValVal	104
Db	8909	GAGGTTGTGATGTGGAGAAATCCTCCGAAGAATCTGAACACAGCCGCATGACATGTC	8968	
OY		105	SerPhehisglnValaAspyrrThrhPheaspArgnaValleuSeraGlueUeuasngu	124
Db	8969	AGCTTTACGAAGGTGAGATTAACCTTTGATACGAATGTGCTCTTAACCTTCGACAG	9028	
OY		125	CyArgrgluWetleuHisglnllelglnArghisluethralalyserHisglyArg	144
Db	9029	TGCAAGAGACCTGGTCCATGAACTGTACAGCGACGACACTTGACGCCCAACCGACGCGC	9088	
OY		145	ValasnaenValPheasaphisPheSerAspCaglugPheUeuAlaleuTyraAsnPro	164
Db	9089	ATTAACCATGTCTTCAACCACTTTCGTGATGTGAATTCCTTTCACTCTCTACAGTCCG	9148	
OY		165	PhegllyanPhylspProHisleuGlnlyLeuCyaspglyleasnylmeUeuasp	184
Db	9149	CACGGAACATCGCACAGCCCAATCTCAAGAGATTTGGACAGGAATCAATAATTGTGAT	9208	
OY		185	GluGUlU 186	
Db	9209	GACAA 9214		
RESULT 39				
ACLI18938/c				
DEFINITION	Mus musculus clone RP24-174O15, WORKING DRAFT SEQUENCE, 8 unordered pieces.			
ACCESSION	ACLI18938			
VERSION	ACLI18938.3 GI:28475940			
KEYWORDS	HMG; HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryotic Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
AUTHORS	Birren,B., Nuebaum,C. and Lander,E.			
JOURNAL TITLE	Mus musculus, Clone RP24-174O15			
REFERENCE AUTHORS	Unpublished 2 (bases 1 to 164258)			
	Birren,B., Lincoln,L., Nuebaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Brown,A., Canatista,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,W., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Glude,S., Gordon,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Komar,A., Karatas,A., Kells,C., Labocque,K., Lamazares,R., Landers,T., Lehoczkay,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,T., McEwan,P., McKernan,K., Meldrim,J., Menues,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,			

```

Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strasman, N., Subramaniam, A., Talamas, J., Teefaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vasiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (21-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164258)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
TITLE
JOURNAL
COMMENT
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 174.O.15
Center clone name: 174.O.15
----- Summary Statistics -----
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162606 bases at least Q40
Consensus quality: 162965 bases at least Q30
Consensus quality: 163183 bases at least Q20
Insert size: 164000; agarose-IP
Insert size: 163558; sum-of-coverage
Quality coverage: 11.0 in Q20 bases; agarose-IP
Quality coverage: 11.0 in Q20 bases; sum-of-coverage
-----
NOTE: This is a 'working draft' sequence. It currently
consists of 8 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1
29707: contig of 29707 bp in length
*
29808 29807: gap of 100 bp
*
29808 32340: contig of 2533 bp in length
*
32341 32440: gap of 100 bp
*
32441 37689: contig of 5249 bp in length
*
37690 37789: gap of 100 bp
*
37790 104529: contig of 66740 bp in length
*
104530 104629: gap of 100 bp

```

```

* 104630 115026: contig of 10397 bp in length
* 115027 115126: gap of 100 bp
* 115127 131677: contig of 16551 bp in length
* 131678 131777: gap of 100 bp
* 131778 158793: contig of 27016 bp in length
* 158794 158893: gap of 100 bp
* 158894 164258: contig of 5365 bp in length.
FEATURES
source
1..164258
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-174015"
/clone_11b="RP24-174015"
/clone_11b="RP24-174015"
1..29707
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
misc_feature
32441..37689
/note="assembly_fragment"
37790..104529
/note="assembly_fragment"
104630..115026
/note="assembly_fragment"
115127..131677
/note="assembly_fragment"
131778..158793
/note="assembly_fragment"
158894..164258
/note="assembly_fragment"
clone_end:T7
vector_side:right"
ORIGIN
Alignment Scores:
Pred. No.: 1.68e-44 Length: 164258
Score: 626.00 Matches: 109
Percent Similarity: 86.26% Conservative: 48
Best Local Similarity: 59.89% Mismatches: 25
Query Match: 65.07% Indels: 0
DB: 2 Gaps: 0
US-10-627-571-2 (1-168) x AC118938 (1-164258)
OY 5 Valpheasnsrlysaenleualvalglalaglnlysllyleuglylmeval 24
DB 87991 GTGTTTATTCAGAAATCTGCGCTTCAAGCCAGAAAGATCTGAGCAATAGACC 87932
OY 25 Serlysserlealatrthrleuileaspasphrsersergluvalleuaggluleu 44
DB 87931 AGCAAAACTGGGCAACATGCTATTGATGACACGACGATGATCTTTGATGACGTG 87872
OY 45 TyrargvalThrarggluYrThrGlnasnlyslgluaglyllysllyleuasn 64
DB 87871 TACAAGTCACGAATAACACACCCACCAACAGAGAGGCCACCAAGATATAGAGAT 87812
OY 65 leuileysrthrvalilelysleualileuetyrargasnanglnpheasnglnasp 84
DB 87811 GCTATCAAGGCGCAATCAAAATGCGATCTCTACCGAACAACAGTTCAGTCAAGAG 87752
OY 85 Gluleualaleumetgluylsrphelyllyllyslvalhlsnglnleualamerthrval 104
DB 87751 GAGGTATTAATGTGGAGAACTCCGAGAGAACTGACCAAGCATGCGATGACATGCTC 87692
OY 105 serphetisglnvalaapYrThrPhaspaparganValleuserargyleuasnqlu 124
DB 87691 AGCTTTCACGAAGTACGATACCTTTCATGACGATGCTCTTACGCTTTCATGACGAG 87632
OY 125 Cysarggluwtelcuhihsnglnlelelglnarghlsleuthralalysserhsglyarg 144
DB 87631 TGCAAGACGCTGTAACATGAACTGTAACAGGACATTTGACACCCAGAACCATGAGCGC 87572

```

Oy		145 ValAsnSerVal ¹ PheArgPHisphSeArSpCYGluPhelLeuAlalaleuTYrAenPro	164
Dd	87571 ATAAACATGTCTTCAACCACACTTTGGTCAGTAGTGGAATTCTTTCCACTCTC TAGGTCGC	87512 	
OY	165 PheGLysNpheLysProHIsleuGlInLyLeuCysAspQlylleAnlysMetLeuasp	184 	
Db	87511 CATGGAACAAGTCGAGGCCCATCTCTCAAGAAGATTGGCAAGAGATCATATAATTGTTAGAC	87452 	
OY	185 GluGlu 186		
Db	87451 GACAAG 87446	*****	
RESULT 40			
LOCUS	ACI07848		
DEFINITION	ACI07848	195574 bp DNA linear HTG 22-MAR-2004	
	Mus musculus chromosome 9 clone RP23-121A19 map 9, *** SEQUENCING		
	IN PROGRESS ***, 8 unordered pieces.		
ACCESSION	ACI07848		
VERSION	ACI07848.5 GI:45589920		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_FULFILLTOP; HTGS_ACTIVEFIN.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 195574) Birren,B., Nusbaum,C. and Lander,E. Mus musculus chromosome 9, clone RP23-321A19 Unpublished 2 (bases 1 to 195574) Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
REFERENCE			
AUTHORS			

TITLE Direct Submission
JOURNAL Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 195574)
AUTHORS Biren, B., Nisbaum, C., Lander, E., Abouelleil, A., Allen, N.,
Anderson, M., Arachchi, H.M., Barn, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boughgalter, B., Camarata, J., Chang, Y.,
Collamore, A., Cook, A., Cooke, P., Corum, B., Dextrallano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fato, S.,
Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Galyon, S.,
Graham, L., Grand-Pierre, N., Haler, N., Hagopian, D., Hago, B.,
Hall, U., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Levine, R.,
Lindblad-Toh, K., Lin, X., Lui, A., Mabbitt, R., Maclean, C.,
McDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Medlind, T., Menees, L., Minova, T., Mienna, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, C., Phunkhang, P., Pierre, N.,
Rechupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Seaman, S., Seeman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.

TITLE	JOURNAL	COMMENT
Submitted (22-MAR-2004) whitehead, Inscitute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Research, 320 Charles Street, Cambridge, MA 02141, USA	On Mar 22, 2004 this sequence version replaced g1:44586703.
Submitted (22-MAR-2004) whitehead, Inscitute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	Research, 320 Charles Street, Cambridge, MA 02141, USA	On Mar 22, 2004 this sequence version replaced g1:44586703.

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L20615
Center clone name: 321_A_19

```
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
```

*	1	7878:	gap of 7878 bp in length
*	7879	7878:	gap of 100 bp
*	7879	28230:	contig of 20252 bp in length
*	28331	28330:	gap of 100 bp
*	28331	59488:	contig of 31158 bp in length
*	59489	59588:	gap of 100 bp
*	59589	100956:	contig of 41368 bp in length
*	100957	101056:	gap of 100 bp
*	101057	101056:	contig of 2478 bp in length
*	101057	103634:	gap of 100 bp
*	103635	103634:	gap of 100 bp
*	103635	142683:	contig of 39049 bp in length
*	142684	142683:	gap of 100 bp
*	142788	154265:	contig of 11480 bp in length
*	154264	154265:	gap of 100 bp
*	154264	154363:	gap of 100 bp
*	154364	195574:	contig of 41211 bp in length

```
FEATURES
        source
            1..195574
                location/Qualifiers
                    /organism="Mus musculus"
                    /mol_type="Genomic DNA"
                    /db_xref="taxon:10090"
                    /chromosome="9"
                    /map="9"
                    /clone="RP23-321A19"
                    /clone_1fb="RPc1-23 Female Mouse BAC"
```

ORIGIN

Alignment Scores:	
Pred. No.:	2, 01e-44
Score:	626.00
Percent Similarity:	86.28
Best Local Similarity:	59.894
Query Match:	65.074
BB:	2
Length:	195574
Matches:	109
Conservative:	48
Mismatches:	25
Indels:	0
Gaps:	0

US-10-627-571-2 (1-188) x AC107848 (1-195574)

```
Dy      5 ValPheAsnSerLysAsnLeuAlaValGlnAlaGlnIlybLybIleuGlYlysMetVal 24
        |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||
Db    16005 GTCTTTAGTTCACAGAATCTGCCTTCACAGCCACAGAAGAAGATCCTGACACAAGTAGCC 16
```

Qy 25 SerLySserIlealaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeu 44
Db 160065 AGCAAAATCTGTGGCCACATGCTGATTGTATGACACCAAGCAGTAGATCTTTGATGAGCTG 160124

45 TyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLysLysIleLysAsp

Df	160125	TATCAAAAGTCAKGAATAATACAKACCCCAACAAGAGAAGGCCCAACAAGATCATGAAGAT	160184
Oy	65	LeuilelyeThrValillelysleualaialeleTyArGaAnaEnglnPheAenglnaaP	84
Df	160185	GCTATTCAGAGTGCGACATCAAAAATTGGCATTCTCTACCGGAAACAACAGTTCAAGTCAAGAG	160244
Oy	85	GluAleualaleuMeGluIylusPheIylusIylusValHlGlInleuaLmetThrValal	104
Df	160245	GAGGTATATATTTGTGAGAGAACTCCGAGAGAAACTAACCCAGCTCGATGACATGATGC	160304
Oy	105	SerPheHlEglValAlAsPyTrThrPheAParGaAnValIleusezArIleuleAuEngln	124
Df	160305	AACCTTCTACGAAGTAGAGTAAGTAATACCTTGTATACAAATGTCTCTCTAAGCTTCGACATGAG	160364
Oy	125	CyArGluMetIeuHlsglnlleleglnarighleuThralalySerHlsgIyarG	144
Df	160365	TGCAGAGACCTGGTATCATGAATCTGATCACAGCGACATTCGACACCAGAACCATGAGACCG	160424
Oy	145	ValAenAenValPheAsphlslPheSerArPySglinPheulaLaalaLeuTyArAnPro	164
Df	160425	ATAAACCAAGTCTTCAACACATTTGCTGATGTGGAATTCCTTTCACATCTCTATATGATCGG	160484
Oy	165	PheGIAsnPhelhyPserProHlsleuglnIylusleCyAsPglYrlleAnaIysmetLeuap	184
Df	160485	CATGGAATCTGCAGGCCCATCTCAAGAGGATTTGCCAAGAGATCAATAAATTTTAC	160544
Oy	185	gluIin 186	
Df	160545	GACAAAG 160550	
RESULT 41			
BC076797		1938 bp mRNA linear VRT 20-SEP-2004	
LOCUS	Xenopus laevis MGC83729 protein, mRNA (CDNA clone MGC:83729		
DEFINITION	IMAGE:6639647), complete cds.		
ACCESSION	BC076797		
VERSION	BC076797.1 GI:49899742		
KEYWORDS	MGC.		
SOURCE ORGANISM	Xenopus laevis (African clawed frog)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;		
REFERENCE AUTHORS	Xenopodinae; Xenopus; Xenopus. 1 (bases 1 to 1938) Klein,S.L., Stransberg,R.L., Wagner,L., Pontius,J., Clifton,S.W. and Richardson,P. Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative Dev. Dyn. 225 (4), 384-391 (2002)		
JOURNAL PUBLISHED REFERENCE AUTHORS	2 (bases 1 to 1938) 12454917 Stransberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shemmen,C.M., Schler,G.D., Altechul,S.F., Zeeberg,B., Buetcow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,D., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L., Scheltz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carimni,P., Prange,C., Rahs,S.S., Loggellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.U., McKernan,K.D., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahay,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Boulford,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skelske,U., Smalls,D.E., Schnerck,A., Schein,J.B., Jones,S.J. and Marra,M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)		
TITLE			
JOURNAL PUBLISHED			
PUBMED	12477932		

REFERENCE 3 (bases 1 to 1938)
AUTHORS Klein,S. and Gerhard,D.S.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA

REMARK NIH-MGC Project
COMMENT Contact: XGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Martha Robbert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liso, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lissa Pirahbu, Parvaneh Saeedi, JR Santos, Angelique Scherch, Ursula Skalska, Duane Smalls, Jeff Sroet, Miranda Tsai, George Yang, Jacquie Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLML at: <http://image.jnl.gov>
Series: IRAK Plate: 158 Row: d Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES
source
1. 1938
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/cdonc="MGC:83729 IMAGE:6639647"
/tissue_type="Oocytes"
/clone_id="NICHG_XGC_OO1"
/lab_host="DH10B"
/notc="vector: pCMV-SPORT6.1"
1. 1938
/gene="WGC83729"
/db_xref="LOCUSID:446389"
94. 702
/gene="WGC83729"
/codon_start=1
/product="WGC83729 protein"
/protein_id="AAH76797.1"
/db_xref="GI:49899743"
/db_xref="LOCUSID:446389"
/translational="MDMDSGDLSRGLSPGPEOFSKSPAVOAKKILSKMATKTMANMLINDRSEIFDELFRKTKRYVKKKKAAHVAIKDYKVAVKGIILRNKQFSSEIJEIVENRRKLNQTCMTAVSFVEVYTPDNNVLSGLLHCKTILHBLVPHLTPKSHSRIDRVFNHFADVEFLTALYSLSEGNYPYLRKICEGVNKLDERVL"

ORIGIN
Alignment Scores:
Pred. No.: 8.55e-45 Length: 1938
Score: 607.00 Matches: 106
Percent Similarity: 83.89% Conservative: 45
Best Local Similarity: 58.89% Mismatches: 29
Query Match: 63.10% Indels: 0
DB: 5 Gaps: 0

US-10-627-571-2 (1-188) x BC076797 (1-1938)

6 PheanSerLySaenLeuAlaValGlnAlGlySerLySileuGlyLySmetValSer 25
151 TTTTGGCTCAAGACCTTGGCTGTTCAAGCCAGAGAGAAATTTCTCAGTAAAGATGGCACA 210
26 TysSerIlealArThrIleuIleasPAspThrSerSerGluValIleuAspGluLeuTyr 45

DB	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	1	Klauser R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Datchenko L., Marusina K., Farmer A.K., Rubin G.M., Hong L., Stedman M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Toobyuk S., Carninci P., Prange C., Raha S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., Mobaw P.J., McKeenan K.J., Malek J.A., Gunaratne P.H., Richards S., Wolley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Viallison D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bonifard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Greenwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzyzinski M.I., Skalska U., Smalton D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	2	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	3	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	4	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	5	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	6	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	7	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	8	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	9	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	10	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	11	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	12	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	13	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	14	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	15	(bases 1 to 1813)	Submitted (02-JUN-2003), National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA	
Db	BC053167	Danio rerio zgc:63960, mRNA (cDNA clone MGC:63960 IMAGE:6791907), complete cds.	BC053167	BC053167.1	GI:31418810	MGC.	Danio rerio (zebrafish)	16			

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cga@bs-re@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.simgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxill.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 117 Row: e Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source	Location/Qualifiers
	1..1813
	/organism="Danio rerio"
	/mol_type="mRNA"
	/db_xref="taxon:7955"
	/clone="MGC:63960 IMAGE:6791907"
	/tissue_type="Kidney, zebrafish"
	/clone_id="NCI CGAP_zkid1"
	/lab_host="DH10B"
	/note="Vector: pCMV-SPORT6.1"
	1..1813
	/gene="zgc:63960"
	/note="synonym: MGC63960"
	/db_xref="LocusID:333322"
	178..786
	/gene="zgc:63960"
	/codon_start=1
	/product="hypothetical protein MGC63960"
	/protein_id="AAH53167.1"
	/db_xref="GI:31418811"
	/db_xref="LocusID:333322"
	/translation="MDSDSGEBSRGSPPHSFNKSIALOAKKILSKATMAVANLTDDTSSEILDELKYASRYTKSKKAHKITKDVTALKIGLYRNHPSPPEMETVERKKRMQNPAMTVVSFYEEYTFPDGIISELLECRDLHELVEHLTWRSGRIDVFHFNFADVDVFTELTGPSBDYRLNLRKI CDGINKLDEBGL"

CDS

gene

ORIGIN

Alignment Scores:

Pred. No.:	3.36e-44	Length:	1813
Score:	600.00	Matches:	106
Best Similarity:	84.44%	Conservative:	46
Best Local Similarity:	58.84%	Mismatches:	28
Query Match:	62.37%	Indels:	0
DB:	5	Gaps:	0

US-10-627-571-2 (1-188) x BC053167 (1-1813)

```

Oy      PheAsnSerLySaenLeuAlaValAGlAlaGlInLySyIleLeuGlyLyMetValSer 25
Db      |||||||.....:|||||||.....:|||||||.....:|||||||.....: 294
Oy      TTCAATTCCAAAAGATTGGCCCTCAGAAGCTCAAAGAAGATTGTGAATAAATGCCACC 294
Db      .....:|||||||.....:|||||||.....:|||||||.....: 294
Oy      LysSerIlealArThrTrtleuIleaspApThrSeSerGiValLeuaspGiuleTy 45
Db      :|||:::|||||||.....:|||||||.....:|||||||.....: 354
Oy      ATGCGCCGTGGCAACTCTTAACAGACGACACCGACGAGATTTCGACGAAGCTCAC 354
Db      .....:|||||||.....:|||||||.....:|||||||.....: 354
Oy      ArgValArhzArgGiUurTrrhngInAsrLyIsGluJalGluLySyIleLyAsnLeu 65
Db      :::::|||||||.....:|||||||.....:|||||||.....: 414
Oy      AAGGCCAGTCGAGAAATACCCAAAGACAAGAGAACCCACAAGATCATCAAGAGATC 414
Db      .....:|||||||.....:|||||||.....:|||||||.....: 414
Oy      IleLySrValIleLySleuAlaIleLeuTyZrgasnaEngInPheasnGlnaspGiU 85
Db      |||:::|||||||.....:|||||||.....:|||||||.....: 474
Oy      ATCAAGATCGCTCGAAGATTGGATTCTTACCGGAACACACAGTTCAAGTCCGAGAG 474
Db      .....:|||||||.....:|||||||.....:|||||||.....: 474
Oy      LeuAlaleuWetGluLySpheLySLyVaValHisGlnleuAlaMetThrValValSer 105

```


where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YVC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep Clone-derived Zebrafish pUC subclones occasionally display inconsistency over the length of mononucleotide A/T runs and conserved TA repeats. Where this is found the longest good quality representation will be submitted.

Repeat names beginning 'Dr' were identified by the Recon repeat discovery system (Zhifeng Bao and Sean Eddy, submitted), and those beginning 'drr' were identified by Rick Waterman (Stephen Johnson lab, WashU). For further information see http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml CH211-12A1 is from a CHORI-211 BAC library

FEATURES

Location/Qualifiers
1. 189797
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-12A1"
/clone_11b="CHORI-211"

ORIGIN

Alignment Scores:
Pred. No.: 4.07e-42 Length: 189797
Score: 600.00 Matches: 106
Percent Similarity: 84.44% Conservative: 46
Best Local Similarity: 58.89% Mismatches: 28
Query Match: 62.37% Indels: 0
DB: 5 Gaps: 0

US-10-627-571-2 (1-188) x BX927313 (1-189797)

QY 6 PheAsnSerIysAsnLeuAlaValAlaGlnAlaGlnLysIleLeuGlyLysMetValSer 25
167674 TTCAAATTCGAAAGATTGGCCCTTCAGGCTCAAAAGAGATTGAGTAAATGGCCACC 167615
QY 26 LysSerIleLeaIhrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTyr 45
167614 ATGCGCGTGGCGAAGCTCTTAACAGACGACACGAGGAGATTCGAGCAAGCTTAC 167555
QY 46 ArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGlnLysLysIleLysAsnLeu 65
167554 AAGGCCAGTCGAGAAATACCAAGACGACAGAGAGAGCCACAGATCATCAAGATGTC 167495
QY 66 TLeuLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnAspGlu 85
167494 ATCAAGATCGCTCGAAGATTCCTTACCCGGAACACACAGTTCAGTCCCTGAGAG 167435
QY 86 LeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAlaMetThrValValSer 105
167434 ATGAGACAGCTGAGCCCTTCAAAAGAGATGACACGCGCCATGACAGCTGAGC 167375
QY 106 PheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125
167374 TTATTAGAAAGTGAGTACATTCGACCGGAGCATTCCTTCAAGAGCTGTTGGAATG 167315
QY 126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgVal 145

DB 167314 AGAGACCTTTCACACAGCTGCTGAGACACCACTTACCTGCGCTCACACGCGGAT 167255

QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTyrAsnProPhe 165
167254 GACCAGCTTTTCAACATTTCCGCGAGATGGATTCCTGACCGAGCTGTAGCGGCATCT 167195

QY 166 GlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeuAspGlu 185
DB 167194 GAAGACTACAGACTTAACCTGAGAGAGATCTCGATGGATTAACAACTCTAGACGAG 167135

RESULT 45

LOCUS AC005339 32360 bp DNA linear PRI 30-JUL-1998
DEFINITION Homo sapiens chromosome 19, cosmid R33729, complete sequence.
ACCESSION AC005339
VERSION AC005339.1 GI:3355454
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 32360)
Lamerdin, J.E., McCready, P.M., Skowronski, E., Adamson, A.W., Burkhardt-Schultz, K., Gordon, L., Kyle, A., Ramirez, M., Stillwagen, S., Phan, H., Velasco, N., Do, L., Regala, W., Terry, A., Ganes, J., Dangnan, L., Poundstone, P., Christensen, M., Georgescu, A., Avila, J., Liu, S., Attix, C., Andreise, T., Trankhim, M., Amico-Keller, G., Coefield, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G., Krommiller, B., Arellano, A., Montgomery, M., Oy, D., Nolan, M., Trom, S., Kobayashi, A., Olsen, A.S. and Carraro, A.V.
Submitted (30-JUL-1998) Joint Genome Institute, Lawrence Livermore National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
Map and sequence oriented from p telomere to centromere. Cosmid R33729 overlaps BAC 48708 to the left from bases 1 to 8,574 and overlaps cosmid R26894 to the right from bases 26,845 to 32,360. Additional chr 19 map and sequence information are available at: <http://www.bio.livl.gov/genome/genome.html>.

Location/Qualifiers
1. 32360
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3 Between D19S883 and MLT1"
/clone="R33729"
/cell_line="5HL2-B"
/clone_11b="UL19NC03 R chromosome 19-specific cosmid library"
/note="Cosmid library constructed at LIVL from flow-sorted chromosomes from human-hamster hybrid 5HL2-B, which carries chromosome 19 as its only human chromosome."

COMMENT

1. 249
/rpt_family="AluXs"
complement(354..693)
/rpt_family="LINE2"
710..787
/rpt_family="MIR"
complement(1403..1521)
/rpt_family="LINE2"
1562..1672
/rpt_family="AluUb"
1689..1780
/rpt_family="GA)n"
complement(2197..2502)
/rpt_family="AluSx"
3589..3609
/rpt_family="GC-rich"

FEATURES

source
1. 32360
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3 Between D19S883 and MLT1"
/clone="R33729"
/cell_line="5HL2-B"
/clone_11b="UL19NC03 R chromosome 19-specific cosmid library"
/note="Cosmid library constructed at LIVL from flow-sorted chromosomes from human-hamster hybrid 5HL2-B, which carries chromosome 19 as its only human chromosome."
1. 249
/rpt_family="AluXs"
complement(354..693)
/rpt_family="LINE2"
710..787
/rpt_family="MIR"
complement(1403..1521)
/rpt_family="LINE2"
1562..1672
/rpt_family="AluUb"
1689..1780
/rpt_family="GA)n"
complement(2197..2502)
/rpt_family="AluSx"
3589..3609
/rpt_family="GC-rich"

```

repeat_region      4266..4572      /rpt_family="AluJc"
repeat_region      4692..4784      /rpt_family="MIR"
repeat_region      4793..4874      /rpt_family="MIR"
repeat_region      /rpt_family="MIR"
repeat_region      complement(4888..5232)
repeat_region      /rpt_family="AluJc"
repeat_region      complement(5239..5545)
repeat_region      /rpt_family="AluSx"
misc_feature       5741..5848
note="predicted exon, program: grai12exons_human_1.3,
frame: 1, quality: good, score: 61.000"
repeat_region      complement(6054..6181)
repeat_region      /rpt_family="FLAM_C"
repeat_region      complement(6188..6262)
repeat_region      /rpt_family="MERSA"
repeat_region      6677..6826
repeat_region      /rpt_family="GC_rich"
repeat_region      complement(6973..7032)
repeat_region      /rpt_family="(CAGA)n"
repeat_region      7764..7895
repeat_region      /rpt_family="MIR"
repeat_region      7919..8022
repeat_region      /rpt_family="LINE2"
repeat_region      8827..8868
repeat_region      /rpt_family="MIR"
repeat_region      9104..9184
repeat_region      /rpt_family="MIR"
repeat_region      9219..9441
repeat_region      /rpt_family="AluJb"
repeat_region      9444..9743
repeat_region      /rpt_family="AluSx"
repeat_region      complement(9773..10218)
repeat_region      /rpt_family="LINE2"
repeat_region      10246..10538
repeat_region      /rpt_family="AluSx"
repeat_region      10911..11179
repeat_region      /rpt_family="AluSx"
repeat_region      11184..11484
repeat_region      /rpt_family="AluSp"
repeat_region      11503..11807
repeat_region      /rpt_family="AluJc"
repeat_region      complement(11856..12145)
repeat_region      /rpt_family="AluSx"
repeat_region      12511..12840
repeat_region      /rpt_family="AluSx"
repeat_region      12868..12968
repeat_region      /rpt_family="AluJ/FLAM"
repeat_region      13135..13228
repeat_region      /rpt_family="LINE2"
repeat_region      complement(13408..13671)
repeat_region      /rpt_family="AluSg"
repeat_region      complement(13674..13880)
repeat_region      /rpt_family="LMB7"
repeat_region      complement(13881..14172)
repeat_region      /rpt_family="AluJ"
repeat_region      complement(14175..14555)
repeat_region      /rpt_family="LMB7"
repeat_region      complement(14560..14854)
repeat_region      /rpt_family="AluSx"
repeat_region      complement(14865..15172)
repeat_region      /rpt_family="AluJc"
repeat_region      15228..15275
repeat_region      /rpt_family="LINE2"
repeat_region      complement(16172..16474)
repeat_region      /rpt_family="AluSg"
repeat_region      complement(16664..16716)
repeat_region      /rpt_family="MIR"
repeat_region      16853..16917
repeat_region      /rpt_family="MIR"
repeat_region      18083..18246
repeat_region      /rpt_family="AluJ"

repeat_region      complement(18247..18276)
repeat_region      /rpt_family="AT_rich"
repeat_region      18551..18586
repeat_region      /rpt_family="AT_rich"
repeat_region      complement(18596..18711)
repeat_region      /rpt_family="FLAM_C"
repeat_region      complement(18725..18991)
repeat_region      /rpt_family="AluSg"
misc_feature       19115..19675
note="predicted exon, program: grai12exons_human_1.3,
frame: 1, quality: excellent, score: 92.000"
repeat_region      20087..20262
repeat_region      /rpt_family="MLTIC"
repeat_region      20269..20558
repeat_region      /rpt_family="AluJ"
repeat_region      20579..20711
repeat_region      /rpt_family="AluJ"
repeat_region      20720..21034
repeat_region      /rpt_family="AluSg"
repeat_region      21049..21201
repeat_region      /rpt_family="AluJ"
repeat_region      21225..21398
repeat_region      /rpt_family="MLTIC"
repeat_region      complement(21329..21549)
repeat_region      /note="BLASTN similarity to Z4385 (1..220) ; match: 0.97,
score: 9.7e-83; database searched: est; H. sapiens partial
cDNA sequence"
misc_feature       21389..21554
note="predicted exon, program: grai12exons_human_1.3,
frame: 1, quality: excellent, score: 77.000"
repeat_region      complement(21587..21852)
repeat_region      /rpt_family="AluJb"
repeat_region      22805..22929
repeat_region      /rpt_family="AluJ"
repeat_region      22943..23243
repeat_region      /rpt_family="AluSx"
repeat_region      23246..23404
repeat_region      /rpt_family="AluJ"
repeat_region      complement(23833..24137)
repeat_region      /rpt_family="AluJc"
repeat_region      24189..24484
repeat_region      /rpt_family="AluJb"
repeat_region      complement(24790..25329)
misc_feature       /note="PDS similarity to overlapping ESTs:
(25329..25201) AA477269 zu43c12.s1 Soares ovary tumor
NBH07 Homo sapiens cDNA clone 740758 3'; (401..273); 100%
identity.-(25329..25201) A477268 zu43c12.r1 Soares ovary
tumor NBH07 Homo sapiens cDNA clone 740758 5'; (221..349);
100% identity.-(25061..24842) A477268 zu43c12.r1 Soares
ovary tumor NBH07 Homo sapiens cDNA clone 740758 5';
(350..570); 98% identity.-(24792..25272) N32339 yw82808.s1
Homo sapiens cDNA clone 258782 3'. Score: 925 identity:
476/480 (99%) .-(25329..25028) AA579149 nt28a04.s1
NCI CGAP Pr1 Homo sapiens cDNA clone IMAGE:915054;
(134..434); 99% identity.-(24790..25061) AA477269
zu43c12.s1 Soares ovary tumor NBH07 Homo sapiens cDNA
clone 740758 3'; (272..1); 100% identity.~Additional EST
matches:
AA581955, AA467935, AI038745, AI041764, T24716"
complement(join(25250..25329,27176..27248,27914..27995,
32121..>32185))
/note="Hypothetical partial human protein"
/codon_start=1
/evidence=not_experimental
/product="R33729_1"
/protein_id="AAC27824.1"
/db_xref="GI:335455"

Alignment Scores:
Pred. No.: 2.77e-42 Length: 32360
Score: 593.00 Matches: 106
Percent Similarity: 79.14% Conservative: 42
Best Local Similarity: 56.68% Mismatches: 39

```

Query Match:	61.64%	Indels:	0
DB:	9	Gaps:	0

6

Gaps:

US-10-627-571-2 (1-188) x AC005339 (1-32360)

QY	2	AlaThrAspVal1PheAsnSerLyAsnLeuAlaValGlnIleGlnIleValSly1LeuGlu	21
Db	19112	GCATGGAACAACCTTCACGACCAACAAAGCTGGCTCTCGAGCGCAGAAAGAGCTCTCGAAT	19171
QY	22	LyMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu	41
Db	19172	AAGATGGCGTCCAGACGACAGTGGTGGCGCTGTGGTGAATGACACAGACGTGAGTGGTGG	19233
QY	42	AspGluLeuValThrArgValThrArgGluTyrThrGlnAsnLySlySglValIleGluSlyS	61
Db	19232	GATAGAGCTGATCCGCGCCACCAAGGAAATTGACGCGCAGCCGCAAGAGAGGCCCAAGAGATG	19291
QY	62	IleLyAsnLeuIleLyThrValIleLeuLeuAlaIleLeuValThrArgAsnAsnIlePhe	81
Db	19292	CTCAAGAACTGGTGCACAGGTGGCGCTGAAAGCTGGAGACTGGCTGGCGGTGGGACCAAGCTG	19351
QY	82	AsnGlnAspGluLeuAlaLeuMetGluLysPheLySlyValValIleGlnLeuAlaMet	101
Db	19352	GGCGATGAGAGCTGGCGCTGTGGTGGCGCTTCGGCACCAGCGCGGCGTGGCTGGCCAGT	19411
QY	102	ThrValValSerPheIleGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu	121
Db	19412	ACGGCCGTCACCTTCCACACAGTGTGACTTCACTTCCACCGGCGGTGGCGCCGCGG	19471
QY	122	LeuAsnGluCysArgGluMetLeuLeuIleGlnIleIleGlnIleArgIleThrAlaLySer	141
Db	19472	CTGCTCGAGTGCACCGACCTGGCTGCACACAGGCGGTGGTGGTCCCACTGACCGCAAGTCC	19531
QY	142	HisGluArgValAsnAsnValPheAspIlePheSerArgPheSerArgPheLeuAlaValLeu	161
Db	19532	CACGGCGGACATCAACCACTGTGGTGGCGACCTGAGCGGACCTTCTGGTGGTGGCTC	19591
QY	162	TyrAsnProPheGluAsnPheLysPheIleLeuGlnIleLysLeuCysArgGluIleAsnIle	181
Db	19592	TACGCGCCCGCCGAGCCTGACCGCTCCCACTCGCGAGATGTGAGAGGCGCTGGGCGG	19651
QY	182	MetLeuAspGluGluIleAsnIle	188
Db	19652	ATGCTGGACAGAGGCGACGCTC	19672

Search completed: July 28, 2005, 20:50:53
Job time : 4523 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 18:13:05 ; Search time 175 Seconds
(without alignments)
1757.825 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962
Sequence: 1 MATDVNSKNIAVQAQKIL.....KPHLQKLCGINKMLDEBN1 188

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh
-Q/cgnt2_1/USPTO.spool.h/US10627571/runat.27027005.154721.27515/app.query.fasta_1.327
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINLEN=0 -MAXLEN=200000000
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=humano.cdi
-LIST=100 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10627571 @CGN1.1.69 @runat.27027005.154721.27515 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgnt2_6/prodata/1/ina/5A.COMB.seq.*
2: /cgnt2_6/prodata/1/ina/5B.COMB.seq.*
3: /cgnt2_6/prodata/1/ina/6A.COMB.seq.*
4: /cgnt2_6/prodata/1/ina/6B.COMB.seq.*
5: /cgnt2_6/prodata/1/ina/PTUS.COMB.seq.*
6: /cgnt2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	527	54.8	1175	4 US-09-799-451-537	Sequence 537, App
2	250	26.0	340	4 US-09-513-999C-14427	Sequence 14427, A
3	166	17.3	252	4 US-09-513-999C-14434	Sequence 14434, A
4	104	10.8	1347	4 US-09-248-796A-6176	Sequence 6176, A
5	96.5	10.0	2229	4 US-09-248-796A-4950	Sequence 4950, Ap
6	92.5	9.6	2253	4 US-09-949-016-3908	Sequence 3908, Ap
7	90.5	9.4	5923	4 US-09-976-594-907	Sequence 907, App
8	90	9.4	1887	4 US-09-601-198-39	Sequence 39, App
9	90	9.4	1664976	4 US-08-916-421B-1	Sequence 1, Appl1
10	90	9.4	1664976	4 US-09-692-570-1	Sequence 1, Appl1
11	89.5	9.3	1142	4 US-09-175-684A-8	Sequence 8, Appl1
12	89	9.3	1975	4 US-09-949-016-975	Sequence 975, App

13	88.5	9.2	1700	4 US-09-016-434-1081	Sequence 1081, Ap
14	88.5	9.2	1700	4 US-09-023-655-898	Sequence 898, App
15	88.5	9.2	1865	4 US-09-270-767-14175	Sequence 14175, A
16	88.5	9.2	3204	4 US-09-710-279-3485	Sequence 3485, Ap
17	88.5	9.2	3627	4 US-09-710-279-4311	Sequence 4311, Ap
18	88.5	9.2	3760	4 US-09-710-279-4029	Sequence 4029, Ap
19	88	9.1	1638	4 US-09-616-289-46	Sequence 46, Appl
20	88	9.1	4206	4 US-09-750-590A-3	Sequence 3, Appl1
21	88	9.1	4637	3 US-08-979-608A-17	Sequence 17, Appl
22	88	9.1	4637	4 US-09-517-849-17	Sequence 17, Appl
23	88	9.1	4697	4 US-09-616-289-17	Sequence 17, Appl
24	88	9.1	4730	4 US-09-750-590A-1	Sequence 1, Appl1
25	87.5	9.1	3172	4 US-09-976-594-940	Sequence 940, App
26	87.5	9.1	3349	3 US-09-336-447A-2	Sequence 2, Appl1
27	87.5	9.1	3349	4 US-09-352-267B-2	Sequence 2, Appl1
28	87.5	9.1	4722	3 US-08-979-608A-14	Sequence 14, Appl
29	87.5	9.1	4722	4 US-09-517-849-14	Sequence 14, Appl
30	87.5	9.1	4722	4 US-09-616-289-14	Sequence 14, Appl
31	87	9.0	4233	3 US-09-541-782-5	Sequence 5, Appl1
32	87	9.0	4233	3 US-09-723-820-5	Sequence 5, Appl1
33	87	9.0	4223	4 US-10-270-085-5	Sequence 5, Appl1
34	86.5	9.0	2469	4 US-09-248-796A-4746	Sequence 4746, Ap
35	86	8.9	1299	4 US-09-248-796A-5857	Sequence 5857, Ap
36	86	8.9	2651	1 US-08-462-949-1	Sequence 1, Appl1
37	86	8.9	2651	1 US-08-023-764B-1	Sequence 1, Appl1
38	85.5	8.9	618	4 US-09-134-000C-1489	Sequence 1489, Ap
39	85.5	8.9	1854	5 PCT-US94-01101-1	Sequence 1, Appl1
40	85.5	8.9	8590	4 US-09-949-016-5562	Sequence 5562, Ap
41	85.5	8.9	10300	4 US-09-949-016-636	Sequence 636, App
42	85	8.8	2637	3 US-09-134-001C-1541	Sequence 1541, Ap
43	84.5	8.8	571	3 US-09-404-879A-84	Sequence 84, Appl
44	84.5	8.8	571	4 US-09-338-933-84	Sequence 84, Appl
45	84.5	8.8	571	4 US-09-215-681-84	Sequence 84, Appl
46	84.5	8.8	571	4 US-09-216-003A-84	Sequence 84, Appl
47	84.5	8.8	571	4 US-09-667-857-84	Sequence 84, Appl
48	84.5	8.8	3025	4 US-09-917-254A-19	Sequence 19, Appl
49	84	8.7	906	4 US-09-248-796A-4219	Sequence 4219, Ap
50	84	8.7	1485	4 US-09-248-796A-53	Sequence 53, Appl
51	84	8.7	58407	4 US-08-916-421B-2	Sequence 2, Appl1
52	84	8.7	58407	4 US-09-692-570-2	Sequence 2, Appl1
53	84	8.7	640681	4 US-09-790-988-1	Sequence 1, Appl1
54	83.5	8.7	734	4 US-08-956-171E-1051	Sequence 1051, Ap
55	83.5	8.7	734	4 US-08-781-986A-1051	Sequence 1051, Ap
56	83.5	8.7	843	4 US-09-134-000C-1958	Sequence 1958, Ap
57	83.5	8.7	1065	4 US-09-175-684A-1	Sequence 1, Appl1
58	83.5	8.7	1088	4 US-09-175-684A-2	Sequence 2, Appl1
59	83.5	8.7	1575	4 US-09-543-681A-615	Sequence 615, App
60	83.5	8.7	1785	4 US-09-117-415B-1	Sequence 1, Appl1
61	83.5	8.7	1862	4 US-09-117-415B-21	Sequence 21, Appl
62	83.5	8.7	1896	4 US-09-117-415B-17	Sequence 17, Appl
63	83.5	8.7	1896	4 US-09-117-415B-19	Sequence 19, Appl
64	83.5	8.7	1950	4 US-09-117-415B-15	Sequence 15, Appl
65	83.5	8.7	2949	4 US-09-023-905A-8	Sequence 8, Appl1
66	83.5	8.7	3446	4 US-09-392-714-13	Sequence 13, Appl
67	83.5	8.7	3636	3 US-09-134-001C-1983	Sequence 1983, Ap
68	83.5	8.7	5954	4 US-09-023-905A-6	Sequence 6, Appl1
69	83	8.6	948	4 US-09-248-796A-2007	Sequence 2007, Ap
70	83	8.6	2553	4 US-09-269-858A-7	Sequence 7, Appl1
71	83	8.6	4853	2 US-08-793-824-1	Sequence 1, Appl1
72	83	8.6	14736	3 US-08-961-527-171	Sequence 171, App
73	83	8.6	164976	4 US-08-916-421B-1	Sequence 1, Appl1
74	83	8.6	164976	4 US-09-692-570-1	Sequence 1, Appl1
75	82.5	8.6	5467	4 US-09-976-594-735	Sequence 735, App
76	82.5	8.6	5967	4 US-09-949-016-1162	Sequence 1162, Ap
77	82.5	8.6	6640	4 US-09-949-016-797	Sequence 797, App
78	82.5	8.6	30549	3 US-09-134-001C-322	Sequence 322, App
79	82	8.5	1026	4 US-09-248-796A-12786	Sequence 12786, A
80	82	8.5	1159	4 US-09-620-312D-56	Sequence 56, Appl
81	82	8.5	1522	4 US-09-248-796A-3452	Sequence 3452, Ap
82	82	8.5	8537	4 US-09-595-684B-30	Sequence 30, Appl
83	82	8.5	8503	4 US-09-620-312D-130	Sequence 130, App
84	81.5	8.5	906	3 US-09-134-001C-78	Sequence 78, Appl
85	81.5	8.5	968	4 US-08-956-171E-553	Sequence 553, App


```
Alignment Scores:
Pred. No.: 9,39e-24 Length: 340
Score: 250.00 Matches: 52
Percent Similarity: 96.36% Conservative: 1
Best Local Similarity: 94.55% Mismatches: 2
Query Match: 25.99% Indels: 0
DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x US-09-513-999C-14427 (1-340)
QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGly 21
DB 174 GCACAGATCTCTTAATTCCTCAAAAACCTGCGCTCAGGACACAAAAGATCTGGGT 233
QY 22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
DB 234 AAATGGTGTCCAAATCCATCGCCACCACTTAATAGACACAAAGTAAGTGGTGGTCTG 293
QY 42 AspGluLeuTyraArgValThrArgGluTyThrGlnAsnLysLys 56
DB 294 GATGAGCTTACAGAGTGACGAGGAGGTACACCAAAAACAAGAG 338

RESULT 3
US-09-513-999C-14434
; Sequence 14434, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclerc, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14434
; LENGTH: 252
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 32
; OTHER INFORMATION: y=c or t
US-09-513-999C-14434

Alignment Scores:
Pred. No.: 6.42e-13 Length: 252
Score: 166.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.26% Indels: 0
DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x US-09-513-999C-14434 (1-252)
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
DB 148 ATGGCCACAGATCTCTTAATTCCTCAAAAACCTGCGCTCAGGACACAAAAGATCTTG 207
QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAsp 35
DB 208 GGTAAATGTGTCTCAAAATCCATCGCCACCACTTAATAGACGAC 252

RESULT 4
US-09-248-796A-6176
; Sequence 6176, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
```

```
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6176
; LENGTH: 1347
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6176

Alignment Scores:
Pred. No.: 0.0011 Length: 1347
Score: 104.00 Matches: 40
Percent Similarity: 42.93% Conservative: 48
Best Local Similarity: 19.51% Mismatches: 51
Query Match: 10.81% Indels: 66
DB: 4 Gaps: 8

US-10-627-571-2 (1-188) x US-09-248-796A-6176 (1-1347)
QY 13 ValGlnAlaGlnLysLysIleLeuGlyLysMetValSerLysSerIleAlaThrThrLeu 32
DB 178 TTACAAAGTATGTAATAATGATTCGAAATTCGAATTAATTCGAATTCGAATTCGATCA 237
QY 33 IleAspAspThrSerSerGluValLeuAspGluLeuTyraArgValThrArgGluTyThr 52
DB 238 ATTGAA-----ACTCAGCTATTGCT 258
QY 53 GlnAsnLys-----LysGlnAlaGlnLysLysIle 62
DB 259 GAAGATTAAGCTTATCATCAAGAAAATTAAAAACAATTAAAGATTAACTAAATAATAT 318
QY 63 LysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnGlnPheAsn 82
DB 319 AATGACTTATTAAGAGAAATTAATTA-----GATCTCAATTAA 363
QY 83 GlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThr 102
DB 364 GAAAAACAACCTTATGATTTATTCGAAACATTAATGAATAATTAAGAAATTAATATCA 423
QY 103 ValValSerPhe-----HisGlnValAspTyThrPheAspArgAsnVal 117
DB 424 AAGGAATCATATGAAGTAAGTTAAATAATTAATGAAGAAACCTTTGATTTTAATTA 483
QY 118-----LeuSerArgLeuLeuAsn 123
DB 484 AATATCGCATCGACAGATCGTACGAGGAGGATTTTCAATTCACACATTCAGATCGAT 543
QY 124 GluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGly 143
DB 544 GAATGAGTGA-----CTTAAAAAAGCTTACAGCAAGAAATTAATCCATCATATGATCT 600
QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeuTyraAsn 163
DB 601 AAAGTTACCAAC-----GAACGTAACTTTTA----- 627
QY 164 ProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeu 183
DB 628-----GAATTACCAACATTAAGATGAATGAATGAATGAATAT 666
QY 184 AspGluGlnAsnIle 188
DB 667 GACGATCAAAACCGTA 681

RESULT 5
US-09-248-796A-4950
```


Db 1026 AAAAAATCA-----TTAGATATTTGAACAATT 1000
Qy 159 A1a1a1eUtyrAsnProhGlyAsnPhelyProhIsuengInlyIsuCySaBpGly 178
Db 999 AATCTTTACTATATTAATTTATTAAT-----GATTACAAATTTATTTACAAGGT 949
Qy 179 11eAsnlyAmec---LeuSpGluGluAAsn 187
Db 948 TTAATATACAATTGCATTAGTAGTGAATAAT 919
RESULT 9
US-08-916-421B-1/c
; Sequence 1, Application US/08916421B
; Patent No. 6503729
; GENERAL INFORMATION:
; APPLICANT: Bult et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6503729
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275
; CURRENT APPLICATION NUMBER: US/08/916,421B
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: US 60/024,428
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (28222)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98266)..(98266)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (98343)..(98343)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (103998)..(103998)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (148948)..(148948)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (163385)..(163385)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191989)..(191989)

OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (191995)..(191995)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (231980)..(231980)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234187)..(234187)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234220)..(234220)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (234814)..(234814)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309398)..(309398)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (309418)..(309418)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312837)..(312837)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (312993)..(312993)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (319226)..(319226)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559167)..(559167)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (559241)..(559241)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (600992)..(600992)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (622708)..(622708)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657081)..(657081)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (657203)..(657203)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (674435)..(674435)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (682442)..(682442)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (713652)..(713652)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (741684)..(741684)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779455)..(779455)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (779676)..(779676)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (855539)..(855539)
; OTHER INFORMATION: n equals a, t, c, or g
; NAME/KEY: misc_feature
; LOCATION: (871619)..(871619)
; OTHER INFORMATION: n equals a, t, c, or g

```

/ NAME/KEY: misc feature
/ LOCATION: (1084830)..(1084830)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1096846)..(1096846)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1119881)..(1119881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1130881)..(1130881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1310988)..(1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1313224)..(1313224)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1345473)..(1345473)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1349491)..(1349491)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1470091)..(1470091)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1569020)..(1569020)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1602912)..(1602912)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1603734)..(1603734)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1637998)..(1637998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ NAME/KEY: misc feature
/ LOCATION: (1664854)..(1664854)
/ OTHER INFORMATION: n equals a, t, c, or g
/ OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1

Alignment Scores:
Pred. No.: 3,42e+03 Length: 1664976
Score: 90.00 Matches: 43
Percent Similarity: 43.85% Conservative: 39
Best Local Similarity: 22.99% Mismatches: 64
Query Match: 9.36% Indels: 41
DB: 4 Gaps: 9

US-10-627-571-2 (1-188) x US-08-916-421B-1 (1-1664976)
QY 9 lyeAsnLeuAlaVal-----GlnAlaGlnLysLys 18
DB 275458 AAAAAGCTTGACATATAAATAAAGATTGGAGATTAATTAAGCTATAATAAAC 275399
QY 19 lIeLeuGlyLysMetValSerLysSerIleAlaThrThrLeuIleAspThrSerSer 38
DB 275398 TTGCTTAATTAAGATGATGATTAATTAAGTCTTAAGTTCAGATTAAGTGAAGT 275339
QY 39 GluVal-----LeuAspGlu-----LeuTyrArgValThrArgGlu----- 50
DB 275338 CAATATCAAAAAGCTTTAGATGAATAATATATATATAAGAAATAATGATCATTAAT 275279
QY 51 ----TyrThrGlnAsnLysLysGluAlaGlnLysLysIleLysAsnLeuIleLysThrVal 69
DB 275278 AAATATATCGAAACATTAAGAAAGAAATGATTAACCTAAAGATTAAGAGTTAAGAGTTA 275219
QY 70 lIeLysLeuAlaIleLeuTyrTyrArgAsnAsnGln-----PheAsn 82
DB 275218 AGTGATATTGCAAAAAATATCATGGAAGAAAGAGAGTATTATGAGCTTTTAAC 275159

```

```

QY 83 GlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThr 102
DB 275158 GAAGATGAATAAATTTATCGATTAATTAATAAAAAACATGACATATTACTCAAAAAAGAA 275099
QY 103 ValValSerPheHisGlnValaLeuPyrThrPheAspArgAsnValLeuSerArgLeuLeu 122
DB 275098 ATTGTTGAATTAATCTGGAATG-----AGCAAAACCAAGGTTTCGAGAAATTATC 275051
QY 123 AsnGluCysArgGluLeuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHis 142
DB 275050 TCCGAATTAAGAGAT-----AGAAATAATATAAGAAA-----GAAAAATA 275009
QY 143 GlyArgValaAsnAsnValPheAspHisPheSerAspCysGlu-PheLeuAlaLeuTyr 162
DB 275008 GGACGAATCAATTAATTA-----ACCTTACAGAAAGAGTAAGAGTTATTA 274961
QY 162 rAsnProPheGlyAsnPhe 168
DB 274960 TAACCATTTAAAAATTTT 274942

RESULT 10
US-09-692-570-1/c
/ Sequence 1, Application US/09692570
/ Patent No. 6797466
/ GENERAL INFORMATION:
/ APPLICANT: Buit et al.
/ TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-
/ Patent No. 6797466
/ TITLE OF INVENTION: jamaaschii
/ FILE REFERENCE: PB275C1
/ CURRENT APPLICATION NUMBER: US/09/692,570
/ PRIOR FILING DATE: 2003-01-14
/ PRIOR APPLICATION NUMBER: US 60/024,428
/ PRIOR FILING DATE: 1996-08-22
/ PRIOR APPLICATION NUMBER: US 08/916,421
/ NUMBER OF SEQ ID NOS: 20
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1664976
/ TYPE: DNA
/ ORGANISM: Methanococcus jamaaschii
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (28222)..(28222)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (28257)..(28258)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (84773)..(84773)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (84808)..(84808)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (84812)..(84812)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (98159)..(98159)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc feature

```

/ LOCATION: (98239)..(98239)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98266)..(98266)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (98343)..(98343)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (103998)..(103998)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (148948)..(148948)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (163385)..(163385)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (191989)..(191989)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (191995)..(191995)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (231980)..(231980)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234187)..(234187)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234220)..(234220)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (234814)..(234814)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309398)..(309398)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (309418)..(309418)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312837)..(312837)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (312993)..(312993)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (319226)..(319226)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559167)..(559167)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (559241)..(559241)

/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (600592)..(600592)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (622708)..(622708)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (657203)..(657203)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (674435)..(674435)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (682442)..(682442)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (713652)..(713652)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (741684)..(741684)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (779455)..(779455)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (779676)..(779676)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (855539)..(855539)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (871619)..(871619)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1084830)..(1084830)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1096846)..(1096846)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1119881)..(1119881)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1130681)..(1130681)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1310988)..(1310988)
/ OTHER INFORMATION: n equals a, t, c, or g
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1313224)..(1313224)
/ OTHER INFORMATION: n equals a, t, c, or g

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1349473)..(1349473)
; OTHER INFORMATION: n equals a, t, c, or g

Alignment Scores:
Pred. No.: 3.42e+03 Length: 1664976
Score: 90.00 Matches: 43
Percent Similarity: 43.85% Conservative: 39
Best Local Similarity: 22.99% Mismatches: 64
Query Match: 9.36% Indels: 41
DB: 4 Gaps: 9

US-10-627-571-2 (1-188) x US-09-692-570-1 (1-1664976)

QY 9 LysAsnLeuAlaVal-----GlnAlaGlnLysLys 18
Db 275458 AAAAATTACCAATAAAATAAAGATTGGAGATAAATTAAGTAAGGCTAATATAAAC 275399
QY 19 IleLeuGlyLysMetValSerLysSerIleAlaThrThrIleuIleAspAspThrSerSer 38
Db 275398 TTGCTAAATGAAGAAGATGATTTATTAAGTCTAAATGAAGATTTCAGAAATGAAGAT 275339
QY 39 GluVal-----LeuAspGlu-----LeuTyraGlyValThrArgGlu----- 50
Db 275338 CAAATACAAAACCTTTAGATGAAATAATATATATATAGAAAATAATGAATCATTAAT 275279
QY 51 ---TyThrGlnAsnLysLysGluAlaGlnLysIleLysAsnLeuIleLysThrVal 69
Db 275278 AATATATCGAAACATTAAAGAAAGAAATGATAGCTAAAGTAAGTAAAGTTAGAGCTTA 275219
QY 70 IleLysLeuAlaIleLeuTyraArgAsnGln-----PheAsn 82
Db 275218 AGTGAATTCGAAAAAATATCATGGAAGAAAAGAGATATTATGAGCTTTTAACT 275159
QY 83 GlnAspGluLeuAlaLeuMetGlnLysPheLysLysValHisGlnLeuAlaMetThr 102
Db 275158 GAAAGATGAAAAAATTATATATCGATTATATATAAACAATGACATATTACTCAAAAAGAA 275099
QY 103 ValIleSerPheHisGlnValAspTyraThrPheAspArgAsnValLeuSerArgLeu 122
Db 275098 ATTGTTGAATTAATCTGGAATG-----AGCAAAACCAAGTTTGCAGAAATTATC 275051
QY 123 AsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHis 142
Db 275050 TCCGAATTAGAGAT-----AGAAAAATATATAGGAAA-----GAAAAATA 275009
QY 143 GlyArgValAsnAsnValPheAspHisPheSerAspCysGlu--PheLeuAlaIleLeuTy 162
Db 275008 GGAAGAAATCAATAAATTA-----ACCTTACAGAGAGAGATGAAGATTATTA 274961
QY 162 AsnProPheGlyAsnPhe 168
Db 274960 TAAACCATTTAAAAATTTT 274942

RESULT 11
US-09-175-684A-8
; Sequence 8, Application US/09175684A
; Patent No. 6593463
; GENERAL INFORMATION:
; APPLICANT: Chen, Li How
; APPLICANT: Meade, Harry M.
; TITLE OF INVENTION: NOVEL MODIFIED MSP-1 NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND METHODS FOR INCREASING MRNA LEVELS AND PROTEIN
; FILE REFERENCE: 10275-133001
; CURRENT APPLICATION NUMBER: US/09/175.684A
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: US 60/085,649
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: US 60/062,592
; PRIOR FILING DATE: 1997-10-20
; NUMBER OF SEQ ID NOS: 19

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 1142
; TYPE: DNA
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1142)
US-09-175-684A-8

Alignment Scores:
Pred. No.: 0.0693 Length: 1142
Score: 89.50 Matches: 50
Percent Similarity: 38.02% Conservative: 23
Best Local Similarity: 26.04% Mismatches: 84
Query Match: 9.30% Indels: 35
DB: 4 Gaps: 7

US-10-627-571-2 (1-188) x US-09-175-684A-8 (1-1142)

QY 7 AsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMetValSerLys 26
Db 487 AACGAGAAAGTACCTGCCCTTCTGAAACAATGAGACCTGTACAAAGCCGCAAGAT 546
QY 27 SerIleAlaThrThrIleuIleAspAspThrSerSerGluValLeuAspGluLeuTyraG 46
Db 547 AAGATTGATCTGTTCTGTATC---CACTGAGGCCAAGGCTCTG----- 588
QY 47 ValThrArgGluTyraGlnAsnLysLysGluAlaGlnLysLysIleLys-----Asn 64
Db 589 -----CAGTACACATATGAGAAAGCAACGTGAGAGTCAAGATCAAGAGCTGAT 639
QY 65 LeuIleLysThrVal---IleLysLeuAlaIleLeuTyraArgAsnGlnPhe----- 81
Db 640 TACCTGAAGACATCCAGATAGCTGCCGATTTCAAGAGAACAACTTCGTCCGA 699
QY 82 -----AsnGlnAspGluLeuAlaLeuMetGlnLysPheLysLys 95
Db 700 ATGCCGATCTGAGCAACGATTACAAACCAACACACTGCTGACCAAGTTC----- 750
QY 96 ValHisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyraThrPheAspArg 115
Db 751 -----CTGAGACCCGAAATGCTCTTCAAAAACCTGGCCCAAG 786
QY 116 AsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArg 135
Db 787 ACCGTCTTGAGCAACTGCTGATGGAACCTGCAAGATGCTGCAATCAGCCAGAC 846
QY 136 HisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCys 155
Db 847 CAGTGTGTAAGAAAGAGTGTCCCAAGACAGCGATGCTTACAGACCTGATGAGAGG 906
QY 156 GluPheLeuAlaIleLeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeu 175
Db 907 GAGGAGTGAAGTGCCTGCTG-----AATCTAAGAGAGAGAGATTAAGTGT 954
QY 176 CysAspGlyIleAsnLysMetLeuAspGluGluAsn 187
Db 955 GTGGAACCCCAATCTACTTGTATACGAGAACAT 990

RESULT 12
US-09-949-016-975
; Sequence 975, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20

```



```

Db      479 ATGCAGCCATGATCAGACCATGACACTCAAGATCCATACAGTATGACACAAAT 538
Qy      117 ValLeuSerArgLeuLeuAenGluCyArgGluMetLeuIsglnIleIleGlnArgHis 136
Db      538 ----- 538
Qy      137 LeuThrAlaLysSerHisGlyArg---ValAsnAsnValPheAspHisPheSerAspCys 155
Db      539 -----AAGGCTCATGCACAAATTAGTTGAGAAATT-----GATGTG 574
Qy      156 GluPheLeuAlaAlaLeuTyraAsnProPheGlyAsnPhelysProHisLeuGlnLys-Le 175
Db      575 GAGAAGGTCTGCTGCTTTGAGAAATCCATATGATGATGCAATTAAGAGTTATATGGAATGAT 634
Qy      175 uCysAspGlyIleAsnLysMetLeuAspGluGluAsnIle 188
Db      635 CCTGGAATCCAGGAATCTATATGATAGACGACGAGAAATATC 674

RESULT 14
US-09-023-655-898
; Sequence 898, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 898:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: G1174071
; US-09-023-655-898

Alignment Scores:
Pred. No.: 0.171 Length: 1700
Score: 88.50 Matches: 41
Percent Similarity: 42.27% Conservative: 41
Best Local Similarity: 21.13% Mismatches: 56
Query Match: 9.20% Indels: 56

```

```

DB:      4 Gaps:      9
US-10-627-571-2 (1-188) x US-09-023-655-898 (1-1700)
Qy      16 GluLysLysIleLeuGlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAsp 35
Db      200 CACGGAGGACGACTTTGGAGAAATGACTCTGAGATCCATCATGCGCTGCTCGAGGAG 259
Qy      36 ThrSerSerGlu-----ValLeuAspGluLeuTyraArgValThrArgGluTyThr 52
Db      260 GAGGACCAAGGAAAGCCCGCGGATCAACAGAGAG-----ATCGACGGCGAGCTCCGC 310
Qy      53 GluAsnLysGluAlaGluLysLysIleLysAsnLeuLe----- 66
Db      311 AGGACAGAGCGGAGCGCCGCCGAGCTCAAGCTGCTGCTGCTCGGACAGAGAGAT 370
Qy      67 -----LysThrValIleLys---LeuAlaIleLeuTyraArgAsnGlnPheAsnGln 83
Db      371 GCGAAGATGACGCTTTATCAAGCAGATGAGAAATCATCATGCGTCAGAGTACTGTGATGAA 430
Qy      84 AspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThrVal 103
Db      431 GATATA-----AGGGCTTCAACCAAGCTGATGATCAGAAACATCTTCACGGCC 478
Qy      104 ValSerPheHis-----GluValAspTyThrPheAspArgAsn 116
Db      479 ATGCAGCCATGATCAGACCATGACACACTCAAGATCCATATGCAAGATATGAGACAAAT 538
Qy      117 ValLeuSerArgLeuLeuAenGluCyArgGluMetLeuIsglnIleIleGlnArgHis 136
Db      538 ----- 538
Qy      137 LeuThrAlaLysSerHisGlyArg---ValAsnAsnValPheAspHisPheSerAspCys 155
Db      539 -----AAGGCTCATGCACAAATTAGTTGAGAAATT-----GATGTG 574
Qy      156 GluPheLeuAlaAlaLeuTyraAsnProPheGlyAsnPhelysProHisLeuGlnLys-Le 175
Db      575 GAGAAGGTCTGCTGCTTTGAGAAATCCATATGATGATGCAATTAAGAGTTATATGGAATGAT 634
Qy      175 uCysAspGlyIleAsnLysMetLeuAspGluGluAsnIle 188
Db      635 CCTGGAATCCAGGAATCTATATGATAGACGACGAGAAATATC 674

RESULT 15
US-09-270-767-14175
; Sequence 14175, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14175
; LENGTH: 1865
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: n means any nucleotide
; US-09-270-767-14175

Alignment Scores:
Pred. No.: 0.196 Length: 1865
Score: 88.50 Matches: 41
Percent Similarity: 38.55% Conservative: 28
Best Local Similarity: 22.91% Mismatches: 67
Query Match: 9.20% Indels: 43
DB:      4 Gaps:      7
US-10-627-571-2 (1-188) x US-09-270-767-14175 (1-1865)

```

```
QY      35 AsptHrSerSerGluValLeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsn 54
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      573 GAGGTACCGCTGAGAGCTAGTGGAGAGCTCAGCTAAAGACGAAAGTTCCTGCGAGCCG 632
QY      55 LysLysGluAlaGluLysLysIleLysAsnLeuIleLysThrValIleLysLeuAlaIle 74
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      633 AATCCAAAGGACGAGGCGCAAAATGGCA--GGGTCAAGGGCAGATCTCGAAAGCTGTCCGGA 689
QY      75 LeuTyrArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGlu-----Lys 91
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      690 CAGGCCAAGTCCAAATCGATCCGCAACCGAGGGCTGCTCCCGGAATGCATGCTGACT 749
QY      92 PheLysLysValIleAsnIle-----LeuAlaMetThrValAlaSerPheHis 107
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      750 TATGGAGAAAGCTCGCGCGAGCAACAGCGTCTTCCGCGAGCGCTCGTCAATTCGGC 809
QY      108 Gln-----ValAspTyrThrPheAspArgAsnValLeuSerArg 120
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      810 GAAAGCGCTGAACAGATGGCCGAGCTCAAGTATTCGTGACGACAAACATCAAGCAGAAC 869
QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      870 TTTTGT-----GAGCCACTGCATCATATGCAGACCAAGACCTCAAGAGAGTA 917
QY      141 SerHisGlyArg----- 144
Db      918 ATGCATCATCGCAAGAGCTCGAGCGCGCGCTAGACTTTGACTTCGAAGCTCCCGCA 977
QY      145 -----ValAsnAsnValPheAspHisPheSerAspCysGluPheLeu 158
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      978 CAGGCCAAGACGATGATTCGTGGTCCGAGAGCAAGTTCGTGAATCGCTCCAGCTG 1037
QY      159 Ala-----AlaLeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeu 175
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1038 GCCCAGGTGGCGATGTTCAATTGCTCGAAGAACATACGAGAGATGTCTCCAGCTG 1094

RESULT 16
US-09-710-279-3485
; Sequence 3485, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3485
; LENGTH: 3204
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-3485

Alignment Scores:
Pred. No.: 0.442 Length: 3204
Score: 88.50 Matches: 46
Percent Similarity: 39.41% Conservative: 34
Best Local Similarity: 22.66% Mismatches: 64
Query Match: 9.20% Indels: 59
DB: 4 Gaps: 9

US-10-627-571-2 (1-188) x US-09-710-279-3485 (1-3204)
QY      23 MetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAsp 42
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2014 TTGATGAGTAAAGAAAGTATGATATCTACAAAGATAGAGATGAGAGAA----- 2064
```

```
QY      43 GluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGluLysLysIle 62
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2065 GAATATTTTCCAGCTACGAGAC----- 2085
QY      63 LysAsnLeuIleLysThrValIleLysLeuAlaIle-----Leu 75
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2086 -----ATAGAGCAGCTCATCGTTTAGTGTTTACTTAGACAAATAGAAATTTG 2136
QY      76 TyrArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLys 95
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2137 GAAAAAGAAAGAAAGAACTTAAAGAAAGAAATATCTCATTTTAGAGAGAGAAATAAAGAG 2196
QY      96 ValHisGlnLeuAlaMetThrValAlaSerPheHisGlnValAspTyrThrPheAspArg 115
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2197 GTTTCATTAATATGTTTAACTTTAAGCTTCTCTATTCAATATGACACACATTTAGAAC 2256
QY      116 AsnValLeuSerArg-----LeuLeuAsnGluCysArgGluMetLeuHisGlnIle 132
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2257 AAGATGATTAATATTTTAAACAATACTTATTTATTAATGAATTAATGATCATCAG--- 2313
QY      133 IleGlnArgHisLeuThrAlaLysSerHisGly----- 143
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2314 CATCGCGCACACACAGAACTAAGCATCATGCATCAATCAAGCCATGCAGGTGATTATGA 2373
QY      144 -----ArgValAsnAsnValPheAspHisPheSerAsp----- 154
Db      2374 AATTAACAACGTTTCTGCATTTTAAATTAATCTTAACGTAATTAATTAATGCGCTGTTTAA 2433
QY      155 -----CysGluPheLeuAlaAlaLeuTyrAsnProPheGlyAsn 167
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2434 GGAGCAAAATGAGACGCTATAGCTGAACAAACAAAGCAGATATCAATAGATGTACC 2493
QY      168 PheLysProHisLeuGlnLys-----LeuCysAsp-----GlyTlleAsnLysMetLeu 183
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2494 GTACATCCCTTGTTCAGAAAGAGCTGTCTATGACTTTTATAGAAATTAACAGAAATTA 2553

QY      184 AspGluGln 186
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2554 GATTAAGAA 2562

RESULT 17
US-09-710-279-4311/c
; Sequence 4311, Application US/09710279
; Patent No. 6703492
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/09/710,279
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4311
; LENGTH: 3627
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-09-710-279-4311

Alignment Scores:
Pred. No.: 0.533 Length: 3627
Score: 88.50 Matches: 46
Percent Similarity: 39.41% Conservative: 34
Best Local Similarity: 22.66% Mismatches: 64
Query Match: 9.20% Indels: 59
DB: 4 Gaps: 9

US-10-627-571-2 (1-188) x US-09-710-279-4311 (1-3627)
```



```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 1638
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (1638)
US-09-616-289-46

Alignment Scores:
Pred. No.: 0.188 Length: 1638
Score: 88.00 Matches: 51
Percent Similarity: 36.28% Conservative: 31
Best Local Similarity: 22.57% Mismatches: 86
Query Match: 9.15% Indels: 58
DB: 4 Gaps: 10

US-10-627-571-2 (1-188) x US-09-616-289-46 (1-1638)

QY 14 GlnAlaGlnLysLysIleuGlyLysMetValSerLysSerIleAlaThr----- 31
Db 619 CAGAGCAAGATGAAGCTCTCCACAGAAAAGCAGAGCCAGCTGTGCAAGAGAACACAC 678
QY 32 LeuIleAspAspThrSerSerGluValLeu-----Asp 42
Db 679 CTGCCGCGGTAGACAGCAAGAGCGCTCTGCGCCGACAGCAAGCTTGAGAGCCATGCCGT 738
QY 43 GlnLeuTyrArgValThrArg-----GlnTyrThrGlnAsnLysGluVal 58
Db 739 GAGCTGCACGCGGACCAACCGCTCTCCAGAGAAAGAGTGTCGACGCGCGCGGAGAGAG 798
QY 59 GlnLysLysIleLysAsnLysLys-----ThrValIleLysLeu 72
Db 799 GAGGAGAAAGCGGACGAGAGGCTGCTGCACCTTCAGGTACACTGAATGACATTCAGCTG 858
QY 73 AlaIleLeuTyr-----ArgAsnAsnGlnPheAsnGlnAspGlnLeuAlaLeuMet 89
Db 859 CAGATGGAACAGACAAATGAGCCCACTCCAAAGCTGCGCCAGAGAAACATGACGTGGCT 918
QY 90 GlnLysPheLysLysValHisGlnLeuAlaMetThrValLysPheHisGlnVal 109
Db 919 GAGAGGCTCAAGAAAGCTGATGAGCATGTGAGCTCGCGAG-----GAGCATATTC 969
QY 110 AspTyrThrPhe-----AspArgAsnValLeuSerArgLeuLeuAsnGln 124
Db 970 GACAAAGCTTCAAAACAAAGCACTTACAAACAGACCTGGATGCCAAGCTCCAGCAG 1029
QY 125 CysArgGlnMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArg 144
Db 1030 GCCCAGGAGATGCTTAAGAGCGCAAGAGCGGCAACGCGGAGAAAGATTTT---CTC 1086
QY 145 ValAsnAsnValPheAspHisPheSerAspCysGlnPheLeu----- 158
Db 1087 CTGAAGAGGACAGTAAGTCCCGACAGAGATGTGTGAGTATGAAGCAGACAGACCAC 1146
QY 159 -----AlaAlaLeuTyrAsnPro----- 164
Db 1147 CTGAAGCAACAGCTTCCCTATTCACAGAGAAAGTTGAGAGAGTTCCAGAACACACTTCC 1206
QY 165 -----PheGlnAsnPheLysProHisLeuGlnLysLeuCysAspGlyIle 179
Db 1207 AAAGCAGAGGAGTATTCAACCACTTCAGCAGGAGATGAGAAAGATGACTAAGAGATC 1266
QY 180 AsnLysMetLeuAspGln 185
Db 1267 AAGAGCTGAGAGAAAGAA 1284

RESULT 20
US-09-750-590A-3
; Sequence 3, Application US/09750590A
; Patent No. 6780987
; GENERAL INFORMATION:

```

```

; APPLICANT: Herman, Ira
; APPLICANT: Welch, Alice
; TITLE OF INVENTION: BETA-CAP73 CONTROL OF NORMAL AND
; TITLE OF INVENTION: ABNORMAL CELL MIGRATION
; FILE REFERENCE: TUI-001CP
; CURRENT APPLICATION NUMBER: US/09/750,590A
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: 60/170,182
; PRIOR FILING DATE: 1998-12-10
; PRIOR APPLICATION NUMBER: 09/733,818
; PRIOR FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 4206
; TYPE: DNA
; ORGANISM: Bos taurus
US-09-750-590A-3

Alignment Scores:
Pred. No.: 0.775 Length: 4206
Score: 88.00 Matches: 36
Percent Similarity: 43.46% Conservative: 47
Best Local Similarity: 18.85% Mismatches: 70
Query Match: 9.15% Indels: 38
DB: 4 Gaps: 7

US-10-627-571-2 (1-188) x US-09-750-590A-3 (1-4206)

QY 9 LysAsnLeuAlaValGlnAlaGlnLysLysIleuGlyLysMetValSerLysSerIle 28
Db 2176 AAAAATGCTCCTTAAAGTAAGTAAGAAATG-----AAAAAGTCA 2217
QY 29 AlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGlnLeuTyrArgValThr 48
Db 2218 CATGATGTAAATGTTGATGATGATTTGAATAAAAGCTTCAGAT-----GTGACA 2265
QY 49 ArgGlnTyrThrGlnAsnLysLysGlnAlaGlnLysLysIle-----LysAsnLeu 65
Db 2266 CACAAATATACGAAAGAAAGAGTTGAAATGCGAAAGTTGTTATGAGAAATGCCAGTTTA 2325
QY 66 IleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnAspGln 85
Db 2326 AGTAATAAATGTGAGCGCGCTGAAACTGTTCATACCTCCGACAGACAGAAAGAA 2385
QY 86 LeuAlaLeuMetGln-----LysPheLysLysValHisGlnLeuAlaMet 101
Db 2386 ATGATGGCTCTGAAATCCAAATATCACTGAACCTTAAGAGCAGCTGTCTGAACCTTAATGA 2445
QY 102 ThrValLysPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
Db 2446 AAATGTGGTGA-----GACCAAGAGAAATATATTCACCTC 2481
QY 122 LeuAsnGlnCysArgGlnMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
Db 2482 ATGCTGAAACAAATGATTTGAAGAAAGACCATGAGTCATGATATGTGCCGTGAAGAAC 2541
QY 142 HisGlnArgValAsnAsnValPheAspHisPheSerAspCysGlnPheLeuAlaLeu 161
Db 2542 CATGAAGATTAATAAATGCC-----TTGAGTGCACA 2574
QY 162 TyrAsnProPheGlnLysPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
Db 2575 TTGATTAATAAACAATAGAGATTAAGTATGATGTAAGAGAAAGTGTGAGATATTAATCA 2634
QY 181 -----LysMetLeuAspGlnLysLeu 188
Db 2635 GAATTTGTGAATAATAAAGATGAGAAACGAATA 2667

RESULT 21
US-08-979-608A-17
; Sequence 17, Application US/08979608A
; Patent No. 6355451

```

Qy		73	AlalleuYr-----ArGanAnslInheanslnarpgluLeuAlLemiet	89
Dd		813	CAGATGGAAACGCACAATGACGCCCACTCCAACTGCCCGAAGAACACTGAGTGA	872
Qy		90	GluUvPheUvlyVtValNlsGlNleuAlamethThyValValSerPheNlsGlNval	109
Dd		873	GAGAGGTCAAAGACTGATTGACGACTATGAGTGAGTCGGCGAG-----GAGCATATC	923
Qy		110	AspTyThrPhe-----AspArgAnValLeuSerAglLeuAnslu	124
Dd		924	GACAAAGCTTTCAACACAAGAAGCTTCAACAGCAGCTGGTGGATGCCAACTCAGCAG	983
Qy		125	CysArgUmetLeuHnlsGlNllelEglaRhlslLeuHrlalylsserHnlsGlyArg	144
Dd		984	GCCCAAGAGTAGTCTTAAGAGGCGAAGAGCGGCGACACCGCGAGAAAGATTT--CTC	1044
Qy		145	ValAsnAnValPheAspHisPheSerAspCysGluPheU-----	158
Dd		1041	CTGAAGAAGCAGTAGTCCCGCAGAGATGTGTAGCTATGATGACGACAGAACACCAC	1100
Qy		159	-----AlalleuTyAsnPro-----	164
Dd		1101	CTGAAGCACAGCTTGCCCTATACACAGAGAAGTTTGAGAGTTCAGAACACACTTCC	1160
Qy		165	-----PheGlyAnPheLyPProHlsleuGlNlyVtLeuCysAserGlyIle	179
Dd		1161	AAAAGCAGCAGGATATTCACCACTTCACAGAGAGATGAAAGAGACTTAAGAGATC	1220
Qy		180	AsnLyMetLeuAspGlu	185
Dd		1221	AAGAGCTGAGAAAGA	1238
<hr/>				
RESULT 22				
US-09-517-849-17				
Sequence 17, Application US/09517849				
Patent No. 6605588				
GENERAL INFORMATION:				
APPLICANT: Lees, Ann M.				
Lees, Robert S.				
Law, Simon W.				
Ajona, Anibal A.				
TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN				
BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND				
TREATING ATHEROSCLEROSIS				
NUMBER OF SEQUENCES: 42				
CORRESPONDENCE ADDRESS:				
ADDRESSEE: Fish & Richardson P.C.				
STREET: 225 Franklin Street				
CITY: Boston				
STATE: MA				
COUNTRY: USA				
ZIP: 02110-2804				
COMPUTER READABLE FORM:				
MEDIUM TYPE: Diskette				
COMPUTER: IBM Compatible				
OPERATING SYSTEM: DOS				
SOFTWARE: FastSeq for Windows Version 2.0				
CURRENT APPLICATION DATA:				
APPLICATION NUMBER: US/09/517,849				
FILING DATE: 02-Mar-2000				
PRIOR APPLICATION DATA:				
APPLICATION NUMBER: 08/979,608				
FILING DATE: 26-Nov-1997				
ATTORNEY/AGENT INFORMATION:				
NAME: Myers, Louis				
REGISTRATION NUMBER: 35,965				
REFERENCE/DOCKET NUMBER: 10797-003001				
TELECOMMUNICATION INFORMATION:				
TELEPHONE: 617/542-5070				
TELEFAX: 617/542-8906				
INFORMATION FOR SEQ ID NO: 17:				
SEQUENCE CHARACTERISTICS:				


```

; LENGTH: 4697 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 3...1592
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-517-849-17

Alignment Scores:
Pred. No.: 0.915 Length: 4697
Score: 88.00 Matches: 51
Percent Similarity: 36.28% Conservative: 31
Best Local Similarity: 22.57% Mismatches: 86
Query Match: 9.15% Indels: 58
DB: 4 Gaps: 10

US-10-627-571-2 (1-188) x US-09-517-849-17 (1-4697)

QY 14 GlnAlaGlnLysIleuGlyMetValSerLysSerIleAlaThr----- 31
DB 573 CAGAGCAGATGATGCTCTTACAGAAAAGCAGAGCCAGCTGTGCAAGAGAACAC 632
QY 32 LeuIleAspAspThrSerSerGluValLeu-----Asp 42
DB 633 CTGCGGGGTGAGCAGCAGCAAGCGCGTCTGCGCCGACAGACTTGAGAGCTTATGCC 692
QY 43 GluLeuTyArgValThrArg-----GluTyThrGlnAsnLysGluAla 58
DB 693 GAGCTCAGCGGACCAACCGCTCCCTCAAGAGAAAGTGTGACGGCGCGGAGAGAG 752
QY 59 GluLysLysIleLysAsnLeuIleLys-----ThrValIleLysLeu 72
DB 753 GAGGAGAGAGGCAAGAGAGTGACCTGCACCTTCAGGTGACCTGAATGACATTCAGCTG 812
QY 73 AlaIleLeuTy-----ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMet 89
DB 813 CAGATGGAAACAGCAGCATATGAGCCCAACTCCAGCTGCGCAAGAAACATGAGCTGGCT 872
QY 90 GluLysPheLysLysValIleIleGlnLeuAlaMetThrValValSerPheHisGlnVal 109
DB 873 GAGAGGCTCAAGAAAGCTGATTGACAGTATGACCTGCGCGAG-----GAGCATATC 923
QY 110 AspTyThrPhe-----AspArgAsnValLeuSerArgLeuLeuAsnGln 124
DB 924 GACAAAGTCTTCAAAACAGAGACCTTACACACAGCAGCTGTGATGCCAAGCTCCAGCAG 983
QY 125 CysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArg 144
DB 984 GCCCAGAGATGCTTAAAGAGGAGGAGCAAGAGCGGACACAGCGGAGAGAAAGATT---CTC 1040
QY 145 ValAsnAsnValPheAspHisPheSerAspCysGluPheLeu----- 158
DB 1041 CTGAAGAGGAGCAGTAAAGTCCAGAGAGATGTGTGAGCTGTGAAGCAGACAGACCAC 1100
QY 159 -----AlaIleLeuTyArgPro----- 164
DB 1101 CTGAGCAACAGCTTGCCTTATACAGAGAGAGTTTGAGAGTTCCAGAACACACTTTC 1160
QY 165 -----PheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIle 179
DB 1161 AAAGAGAGGAGATTTTCAACCCACATTCACAGCAGAGATGAGAAAGATGACTAAGAAATC 1220
QY 180 AsnLysMetLeuAspGlu 185
DB 1221 AAGAACTGTGAGAAAGAA 1238

RESULT 23
US-09-616-289-17
; Sequence 17, Application US/09616289
; Patent No. 6632923
; GENERAL INFORMATION:

```

```

; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; PRIORITY FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 4697
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)...(1592)
US-09-616-289-17

Alignment Scores:
Pred. No.: 0.915 Length: 4697
Score: 88.00 Matches: 51
Percent Similarity: 36.28% Conservative: 31
Best Local Similarity: 22.57% Mismatches: 86
Query Match: 9.15% Indels: 58
DB: 4 Gaps: 10

US-10-627-571-2 (1-188) x US-09-616-289-17 (1-4697)

QY 14 GlnAlaGlnLysIleuGlyMetValSerLysSerIleAlaThr----- 31
DB 573 CAGAGCAGATGATGCTCTTACAGAAAAGCAGAGCCAGCTGTGCAAGAGAACAC 632
QY 32 LeuIleAspAspThrSerSerGluValLeu-----Asp 42
DB 633 CTGCGGGGTGAGCAGCAGCAAGCGCGTCTGCGCCGACAGACTTGAGAGCTTATGCC 692
QY 43 GluLeuTyArgValThrArg-----GluTyThrGlnAsnLysGluAla 58
DB 693 GAGCTCAGCGGACCAACCGCTCCCTCAAGAGAAAGTGTGACGGCGCGGAGAGAG 752
QY 59 GluLysLysIleLysAsnLeuIleLys-----ThrValIleLysLeu 72
DB 753 GAGGAGAGAGGCAAGAGAGTGACCTGCACCTTCAGGTGACCTGAATGACATTCAGCTG 812
QY 73 AlaIleLeuTy-----ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMet 89
DB 813 CAGATGGAAACAGCAGCATATGAGCCCAACTCCAGCTGCGCAAGAAACATGAGCTGGCT 872
QY 90 GluLysPheLysLysValIleIleGlnLeuAlaMetThrValValSerPheHisGlnVal 109
DB 873 GAGAGGCTCAAGAAAGCTGATTGACAGTATGACCTGCGCGAG-----GAGCATATC 923
QY 110 AspTyThrPhe-----AspArgAsnValLeuSerArgLeuLeuAsnGln 124
DB 924 GACAAAGTCTTCAAAACAGAGACCTTACACACAGCAGCTGTGATGCCAAGCTCCAGCAG 983
QY 125 CysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArg 144
DB 984 GCCCAGAGATGCTTAAAGAGGAGGAGCAAGAGCGGACACAGCGGAGAGAAAGATT---CTC 1040
QY 145 ValAsnAsnValPheAspHisPheSerAspCysGluPheLeu----- 158
DB 1041 CTGAAGAGGAGCAGTAAAGTCCAGAGAGATGTGTGAGCTGTGAAGCAGACAGACCAC 1100
QY 159 -----AlaIleLeuTyArgPro----- 164
DB 1101 CTGAGCAACAGCTTGCCTTATACAGAGAGAGTTTGAGAGTTCCAGAACACACTTTC 1160
QY 165 -----PheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIle 179
DB 1161 AAAGAGAGGAGATTTTCAACCCACATTCACAGCAGAGATGAGAAAGATGACTAAGAAATC 1220
QY 180 AsnLysMetLeuAspGlu 185
DB 1221 AAGAACTGTGAGAAAGAA 1238

```

Db 1041 CTGAAGAGCAGTAGAGTCCACAGAGATGTGTAGCTATGAACGACAAAGACCAC 1100
Qy 159 -----AlaAlaLeuTYrAsnPro----- 164
Db 1101 CTGAAGCAGACGCTTGCCTATACACAGAGATTGTAGAGATTCAGAAACACTTTC 1160
Qy 165 -----PheGlyAsnPhelyProHisLeuGlnIleuGlySerAspGlyIle 179
Db 1161 AAAAGCAGCAGATATTCACCACTTCAGAGAGATGAGAAAGATGACTTAAGAGATC 1220
Qy 180 AsnIleuMetLeuAspGlu 185
Db 1221 AAGAGCTGAGAAAGAA 1238

RESULT 24
US-09-750-590A-1
Sequence 1, Application US/09750590A
Patent No. 6780987
GENERAL INFORMATION:
APPLICANT: Herman, Ira
APPLICANT: Welch, Alice
TITLE OF INVENTION: BETA-CAP73 CONTROL OF NORMAL AND
TITLE OF INVENTION: ABNORMAL CELL MIGRATION
FILE REFERENCE: TUI-001CP
CURRENT APPLICATION NUMBER: US/09/750,590A
CURRENT FILING DATE: 2000-12-28
PRIOR APPLICATION NUMBER: 66/170,182
PRIOR FILING DATE: 1999-12-10
PRIOR APPLICATION NUMBER: 09/733,818
PRIOR FILING DATE: 2000-12-08
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 4730
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: CDS
LOCATION: (392)...(4597)
US-09-750-590A-1

Alignment Scores:
Pred. No.: 0.925 Length: 4730
Score: 88.00 Matches: 36
Percent Similarity: 43.46% Conservative: 47
Best Local Similarity: 18.85% Mismatches: 70
Query Match: 9.15% Indels: 38
DB: 4 Gaps: 7

US-10-627-571-2 (1-188) x US-09-750-590A-1 (1-4730)

Qy 9 LysAsnLeuAlaValGlnAlaGlnIleuGlyLeuGlyMetValSerIle 28
Db 2567 AAAAATGTCCTTAAAGTAAGTGAAGAAATG-----AAAAGTCA 2608
Qy 29 AlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTYrArgValThr 48
Db 2609 CATGATGATATGTGTGATGATTTGAATAAAGCTTCAGAT-----GTGACA 2656
Qy 49 ArgGluTYrThrGlnIleuAsnIleuGlyValGlnIleuGlyLeu-----LysAsnLeu 65
Db 2657 CACAAATATACAAAGAAAGATTGAAATGGAAGATGCTTATGTGAAAATGCCAGTTTA 2716
Qy 66 IleLeuThrValIleLeuLeuAlaIleLeuTYrArgAsnAsnGlnPheAsnGlnAspGlu 85
Db 2717 AGRAAAATATCTACGCCGCTGAAACTGTTCATTCCTCCGAGAGACAGCAAAAAGAA 2776
Qy 86 LeuAlaLeuMetGlu-----LysPheIleLeuIleValHisGlnLeuAlaMet 101
Db 2777 ATGATGCTCTGAATTCATATATACCTGAACTTAAGAGACGCTGCTGAACTTAATTA 2836
Qy 102 ThrValValSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArgLeu 121

Db 2637 AATATGTGTGA-----GACCAAGAGAAAATATATTCATCTC 2872
Qy 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIleSer 141
Db 2873 ATGCTGAAAACAAATGATTTGAAAAAGACCATGATGATCAGATATGCCCCGTGAAAACC 2932
Qy 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeu 161
Db 2933 CATGAAGAGATTAAACTGCC-----TTGAGTGCACA 2965
Qy 162 TYrAsnProPheGlyAsnPhelyProHisLeuGlnIleuGlySerAspGlyIleAsn--- 180
Db 2966 TTGATTAACCAATGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAATCA 3025
Qy 181 -----LysMetLeuAspGluIleAsnIle 188
Db 3026 GAATTTGTGAATAATAAAGATGAGAACGAAATTA 3058

RESULT 25
US-09-976-594-940
Sequence 940, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Fumess, Michael
APPLICANT: Buchinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT APPLICATION NUMBER: US/09/976,594
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 940
LENGTH: 3172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6673549 481154.3
US-09-976-594-940

Alignment Scores:
Pred. No.: 0.59 Length: 3172
Score: 87.50 Matches: 44
Percent Similarity: 42.18% Conservative: 45
Best Local Similarity: 20.85% Mismatches: 87
Query Match: 9.10% Indels: 35
DB: 4 Gaps: 8

US-10-627-571-2 (1-188) x US-09-976-594-940 (1-3172)

Qy 9 LysAsnLeuAlaValGlnAlaGlnIleuGlyLeuGlyMetValSerIle 28
Db 1140 AGGAATTTG-----GCCCAACGAAATTAATTAAGAAATGAGTCTTAAGTAA 1190
Qy 29 AlaThrThrLeuIleAspAspThr-----SerSerGluValLeuAspGluLeu 44
Db 1191 GAACAACAATTCGACAGAAACAAAGCTTTAAAGACACAGAAACAATGAAAGAGCTTA 1250
Qy 45 TYrArgValThrArgGluTYrThrGlnAsnIleuGlyValGlnIleuGlyIleu 63
Db 1251 GTATGCAACCTTCTCCGATGACTCAATCAAAATTAAGAAAGAGAAAGCAAAAGTCAAG 1310
Qy 64 AsnLeuIleLeuThrValIleLeuLeuAlaIleLeuTYrArgAsnAsnGlnPheAsnGln 83
Db 1311 GATTTCTGAAGCTCAGCAAAATATACCAACATTTGTTAAAGAAATGAAAGCAAGAT 1370
Qy 84 AspGluLeuAlaLeuMetGluIleuPheIleuIleValHisGlnIleuAlaMetThrVal 103
Db 1371 CTTGAATCAGATATACAAAGAAAGAAATGAAATTTATTCGAGAGCTGAGAGATTT 1430
Qy 104 ValSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArgLeuLeuAsn 123

```

Db      1431 GCTAACTGATGACAC---ATTGGAATGAAAGAAATTTGTTAACTTACTCCAC 1487
      124 GluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeu----- 137
      1488 AAAAGCTCATGAAAGAAATGAAATGAAAGAAAGGATGAAATGCTTAATAATGAACCTT 1547
      138 -----ThralaIalysSerHisGlyArg-----ValAsnAsnValPheAsp 150
      1548 GAAATTCGAGAAATAGTGTGCGCTTACTGACAGAAAGAAAGCTACAAATTCATGCTGAAA 1607
      151 HisPheSerAspCysGluPheLeuAlaIleuTyraAsnProPheGlyAsnPhelysPro 170
      1608 CACGCCAACATGTTACATCAGAGAGACATCAAAACGATGTGCGCAAAATTTGTATCA 1667
      171 HisLeuGlnIlyLeu-----CysAspGlyIleAsnIys----- 181
      1668 AAACCTTCAGGAATGAAAGAAAGAAAGAAAGCCAGTTAAATTAACATTGACAGACTTGCC 1727
      182 -----MetLeuAspGluGluAsnIle 188
      1728 AACACGATCAATGATGACAGAGAGATGCTG 1760

RESULT 26
US-09-336-447A-2
; Sequence 2, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP21 AND USP22 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3349
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-09-336-447A-2

Alignment Scores:
Pred. No.: 0.64 Length: 3349
Score: 87.50 Matches: 47
Percent Similarity: 40.36% Conservative: 43
Best Local Similarity: 21.08% Mismatches: 66
Query Match: 9.10% Indels: 67
DB: Gaps: 10

US-10-627-571-2 (1-188) x US-09-336-447A-2 (1-3349)
      4 AspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlyAsnIleLeuGlyIysMet 23
      1296 GACATTCCTTAACACCAAGATGATCAAGATCTTCAGAGAGGATGAAGGT----- 1349
      24 ValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGlyVal----- 40
      1350 -----CTGTGATATGAGGTGGGTGATTAACCGAGAC 1382
      41 LeuAspGluLeuTyraGlyValThrArgGluTyThrGlnAsnIlySerGluAlaGluIys 60
      1383 ATTAATTCATCTTCATGATCTTACTGACAAACCAAGATGACATCAAAAGGCTTAAGAGG 1442
      61 LysIleIysAsnLeuIleIysThrValIleIysLeuAla-----IleLeuTy 76
      1443 GGGGTAAAGAGCTTGAATTAATGAGGTGGGTGATTAAGCCGAGACATTAATTCATTCAT 1502
      77 ArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeu----- 88

```

```

Db      1503 GATGATGTTGCTGACACCAAGATGACATTTGTAATAAACAAAGCTGACATCAAGGCTTT 1562
      89 -----MetGluIlyPheIlySerIysValIHisGlnLeuAlaMetThrValIysSer 105
      1563 AATAAGAGAGTGAAGAGAGCTTGATTAAGAGAGGTGGGTGATTAAGCCGAGACATTTGGTTCA 1622
      106 PheHis-----GlnValAspTy 111
      1623 CTTCATGATGATGTTGCCACCAAGCTGACATTTGCTTAATAAACCAAGCGGATATCAAA 1682
      112 ThrPheAspArgAsnValIleuSerArgLeuIleuAsnGluCysArgGluMetLeuHisGln 131
      1683 ACACCTGAAGAAACATGTGCAAGAGAAATTAATAATCTTAAGCGGTGCGCTGCTGATCAG 1742
      132 IleIleGlnArgHisLeuThrAlaIysSerHisGlyArgValAsnAsnValPheAsp--- 150
      1743 -----AAAGCGGATATTTGATTAATAACATCAACAAATTAATGATGAGCTG 1784
      151 -----HisPheSerAspCysGluPheLeuAlaIleuTyraAsnPro 164
      1785 GCACACACGACAGATCAGATAGCTCTGATACAAACACTT----- 1826
      165 PheGlyAsnPhelysProHisLeuGlnIly--LeuCysAspGlyIleAsnIlyMetLeu 183
      1827 -----AAAAACAATGTCGAGAGAGGTTATTGATCTTAAGCGGTGCTCATTT 1874
      184 AspGluGlu 186
      1875 GATCAAAA 1883

RESULT 27
US-09-952-267B-2
; Sequence 2, Application US/0995267B
; Patent No. 6753417
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP21 AND USP22 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/952,267B
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US/09/336,447
; PRIOR FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3349
; TYPE: DNA
; ORGANISM: Moraxella catarrhalis
US-09-952-267B-2

Alignment Scores:
Pred. No.: 0.64 Length: 3349
Score: 87.50 Matches: 47
Percent Similarity: 40.36% Conservative: 43
Best Local Similarity: 21.08% Mismatches: 66
Query Match: 9.10% Indels: 67
DB: Gaps: 10

US-10-627-571-2 (1-188) x US-09-952-267B-2 (1-3349)
      4 AspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlyIleuGlyIysMet 23
      1296 GACATTCCTTAACACCAAGATGATCAAGATCTTCAGAGAGGATGAAGGT----- 1349
      24 ValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGlyVal----- 40
      1350 -----CTGTGATATGAGGTGGGTGATTAACCGAGAC 1382

```



```

RESULT 29
US-09-517-849-14
: Sequence 14, Application US/09517849
: Patent No. 6605368
: GENERAL INFORMATION:
: APPLICANT: Lees, Ann M.
:               Lees, Robert S.
:               Law, Simon W.
:               Arjona, Anibal A.
: TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN
:                   BINDING PROTEINS AND THEIR USES IN DIAGNOSING AND
:                   TREATING ATHEROSCLEROSIS
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/517,849
: FILING DATE: 02-Mar-2000
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/979,608
: FILING DATE: 26-NOV-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Myers, Louis
: REGISTRATION NUMBER: 35,965
: REFERENCE/DOCKET NUMBER: 10797-003001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: INFORMATION FOR SEQ ID NO: 14:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4722 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: Coding Sequence
: LOCATION: 61..1731
: SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-517-849-14

Alignment Scores:
Pred. No.:          1.07          Length:          4722
Score:              87.50          Matches:         42
Percent Similarity: 36.79%         Conservative:    36
Best Local Similarity: 19.81%      Mismatches:     91
Query Match:        9.10%          Indels:         43
DB:                  4             Gaps:            7

US-10-627-571-2 (1-188) x US-09-517-849-14 (1-4722)
Oy      13  ValGlnAlaGlnLysLysIleLeuGlyLysMetValSerLysSerIleAlaThrLLeu 32
Db      721  GTCGAGGAGAAAGACACCTGCTGTCGCGAGCAACAGAGCCATCTGCGCCGACAGAG 780
Oy      33  IleApApPThrSerSerGluValIleuApGluLeuTyArgValThrArgGluTyThr 52
Db      781  CTCGAGAGCCTTGCCGCGAGGCTGCGACGCGCAACACCGCTGCTCAAGAAAGAGTG 840
Oy      53  GlnAenLysLysGlnAlaGluLysLysIleLysAsnLennLlys----- 67
Db      841  CAGCGAGCCCGAGAGGAGGAGAAAGCCGACAGAGGTGACGTCACTTCACAGATGACG 900

```

```

QY 68 -----ThrValIleuValleIleuThr-----ArgAsnAngInPheAsnGln 83
Db 901 CTCAACGACACATTCACGCTGACAGATGAGACGACACACAGCGGCAATCTCCAGCTGGCCG 960
QY 84 AspGluLeuAlaLeuMetGluLysPheIlyIlyValHisGlnLeuAlaMetThrVal 103
Db 961 GAGAAACATGGAGCTGGCCGACGGGCTCAAGAAAGCTGATTGAGACAGTACAGACTGGACGAA 1020
QY 104 ValSerPheHisGlnValAAspTyrThrPhe-----AspArgAsnValLeu 118
Db 1021 -----GAGCACATCTGACAAAGCTCTCAACACAGATCTGCAGCAGCACTGGTG 1071
QY 119 SerArgLeuLeuAAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThr 138
Db 1072 GACGCCAAGCTCCACAGACGCCGACAGATGCTGAAGAAGGACGAGACGCGGACACAGCGG 1131
QY 139 AlaIysSerHisGlyArgValAlaAsnAsnValPheAspHisPheSerAspCyGluPheLeu 158
Db 1132 GAGAAAGGACTTT---CTCTCGAAGGAGGCGCTGGAGTCCGACAGATGTGGAGCTGATG 1188
QY 159 -----AlaAlaLeuTyrKaaPro----- 164
Db 1189 AAGCAACGAGACCCACCTGAGACGACGACTTGGTATCACACAGACAACTTTGAGAGG 1248
QY 165 -----PheGluAAsnPheLysProHisLeuGln 173
Db 1249 TTCCGAAACACTTTTCCAAAGACGAGCGAGGTGTTTCAACACATTCAAACAGAAATGGAA 1308
QY 174 LysLeuCyAspArgGlyIleAsnIlyMetLeuAAspGlu 185
Db 1309 AAGATGACAAAGAAAGATTCAGAACCTCGAGAAAG 1344

```

```

RESULT 30
US-09-616-289-14
; Sequence 14, Application US/09616289
; Patent No. 6632523
; GENERAL INFORMATION:
; APPLICANT: Lees, Ann M.
; APPLICANT: Lees, Robert S.
; APPLICANT: Law, Simon W.
; APPLICANT: Arjona, Anibal A.
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING
; TITLE OF INVENTION: PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING
; TITLE OF INVENTION: ATHEROSCLEROSIS
; FILE REFERENCE: 10797-004001
; CURRENT APPLICATION NUMBER: US/09/616,289
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 4722
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)...(1731)
US-09-616-289-14

Alignment Scores:
Pred. No.: 1.07 Length: 4722
Score: 87.50 Matches: 42
Percent Similarity: 36.79% Conservative: 36
Best Local Similarity: 19.81% Mismatches: 91
Query Match: 9.10% Indels: 43
DB: 4 Gaps: 7

```

[illegible]

DB:	3	Gaps:	9
US-10-627-571-2 (1-168) x US-09-541-782-5 (1-4223)			
QY	9	LysaenleuAvalGlnaGlnlybyleuGlylybweValSerlybSerile	28
DB	2885	CAGATATATGACATGCTCTCCACAGAG-----TTGGAAATAGGAATTTCTTCGGAATTG	2938
QY	29	AlaThrThrleuIleAraAraPThrSerSer-----GluValIleuAraGlu	43
DB	2939	ATGAAATTGCAGAAAGACATGAAAGAAATCATACCGACACACTTGTTCAGAAATTAAGATCA	2998
QY	44	LeuTyraArgValThrArgGluTyThrGlaAsnlybyleGlnaGluTybyleyle	63
DB	2999	CTATATATATTTACACGACTACTCATGAAATATCAACGAAAGAGCTAATGTATGTGTCGCT	3058
QY	64	AsnleuIleuTybThrValIleTybLeuValIleleuTyraGlnaGlnGlnPheAsnGln	83
DB	3059	AACGATTTGATGCTGCTCGTTAAACCTGC-----ACACATCTTTGAACGAT	3106
QY	84	AspGluLeuAlaLeuMetGlu-----LysPheTyblybyleVal	96
DB	3107	GCAGATATATATTAAGTGAATTCATATCATATCATCAAAATCCAAATTTGAATCCAGCA	3166
QY	97	HisGlnleuAlaMetThrValValSerPheHisGlnValAspTyThrPheAraGln	116
DB	3167	CAAGATTTGATTTGCTAATATT-----GGTAA	3193
QY	117	ValLeuSerArgleuLeuAenglyCyArgGluMetLeuHis-----GlnIleile	133
DB	3194	ATTGTTTCAATTTTTCACGACACAGCAAACTTTGTATACCAAGCGGATATCTTA	3253
QY	134	GlnArgHisleu-----ThrAlaTybSerHisGlyArgValAsnAraPheAraPheHis	151
DB	3254	CATTCACATCTCATGATATCAAACTGCAATATATAGGAAGAACAAACGAATATGAAATAC	3313
QY	152	PheSerArgCySGluPheleu-----AlaAlaLeuTyraSnProbegIlyAsnPhe	168
DB	3314	CGTTTCAGAA---GAGTTTTTACGGAATGCGCGCTTCCACAGCAAGAAATCGTGCGTCCAA	3370
QY	169	LysPheHisleuGlnlyblybLeuCybAraPglyIleAsnlybMetLeuAraPgluGlna	187
DB	3371	AAAGAAAGCATTCAAAAGACAGTTGAAATATGATCTCACTGCTTGCAGTAAAGC	3427
RESULT 32			
US-09-723-820-5			
; Sequence 5, Application US/09723820			
; Patent No. 6468760			
; GENERAL INFORMATION:			
; APPLICANT: Nislow, Corey			
; APPLICANT: Sakowicz, Roman			
; APPLICANT: Beraud, Christophe			
; TITLE OR INVENTION: Antifungal Assay			
; FILE REFERENCE: 1015			
; CURRENT APPLICATION NUMBER: US/09/723,820			
; CURRENT FILING DATE: 2000-11-28			
; PRIOR APPLICATION NUMBER: 09/541,782			
; PRIOR FILING DATE: 2000-04-03			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; LENGTH: 4223			
; TYPE: DNA			
; ORGANISM: S.pombe			
US-09-723-820-5			
Alignment Scores:			
Pred. No.:	1.05	Length:	4223
Score:	87.00	Matches:	42
Percent Similarity:	41.71%	Conservative:	42
Best Local Similarity:	21.11%	Mismatches:	78
Query Match:	9.04%	Indels:	38
	3	Gaps:	9


```
TELEFAX: 201-753-6237
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2651 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-462-949-1

Alignment Scores:
Pred. No.: 0.708 Length: 2651
Score: 86.00 Matches: 45
Percent Similarity: 34.53% Conservative: 32
Best Local Similarity: 20.18% Mismatches: 50
Query Match: 8.94% Indels: 96
DB: 1 Gaps: 10

US-10-627-571-2 (1-188) x US-08-462-949-1 (1-2651)
QY 31 ThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTyraArgValThrArgGlu 50
   |||||
DB 1768 ACTCTCCGCCCGGATGATGCTCCCTTAATGATCTTAATCTTTTTCACATTAAGCCAGTCC 1709
   |||||
QY 51 TyrThrGln-----AsnLysLysGluAlaGluLysLysIleLys 63
   |||||
DB 1708 CAGACCGGAACCTTTTACATCGTCAGATTACCTGTATGACACGCGATGAGAACTTATCAAA 1649
   |||||
QY 64 AsnLysLysLysThrValIle-----LysLeuAlaIleLeuTyraArgAsnAsnGln 80
   |||||
DB 1648 CACCTTCTTCAGGTTTCTTTTAAATACCAATACGTTGCTCTGTATCGAATATCATACG 1589
   |||||
QY 81 Phe-----AsnGlnAspGluLeuAlaLeuMetGlu 90
   |||||
DB 1588 TTTCACGGGTTCCGTTCCAGGTTCTACGTCGACACGAGGCTTCTTCGGGCTTCTT--- 1532
   |||||
QY 91 LysPheLysLysLysValHisGlnLeuAla----- 100
   |||||
DB 1531 GATATTTCACCGGTTATCCATCAGTTGAATATCACTTGTCATATGCAATTTCGTCGCC 1472
   |||||
QY 101 -----MetThrValIleSerPheHis 107
   |||||
DB 1471 AATATTAACAGAGATTGGTAGCCTCAAGTTCCATGTATATTACATTATAGGCTTCAC 1412
   |||||
QY 107 ----- 107
   |||||
DB 1411 CTTGAGAGGAAAGATATTGATAACCCGGAATCAACTCATTTGGCATCGAGTTCTTTCAT 1352
   |||||
QY 108 -----GlnValAspTyrThrPheAspArgAsnValLeuSerArg--- 120
   |||||
DB 1351 CTTCACTGTGTTGCTTCTGTCGAAACATGACATCTGAAGAATCTTCTCCACCTGGAA 1292
   |||||
QY 121 -----LeuLeuAsnGluCysArgGluMetLeuHisGlnIle----- 132
   |||||
DB 1291 TCTCAACCGCTTGCTTTCATCATTAATGACTCCCGATATATGCTGGAACATCTGCGGTGA 1232
   |||||
QY 133 -----LysGlnArgHisLeuThrAlaLysSerHisGlyArg----- 144
   |||||
DB 1231 TTTCGGAATGCGGATCTTTCAGCATCTGCGGCGAAGCAGATGGTAGATATCGGTGT 1172
   |||||
QY 145 -----ValAsnAsnValPheAspHisPheSerArgPheCysGluPhe 157
   |||||
DB 1171 CTTGAACCTGTGTCATATTGTGTGAATAAGTCAATCTTCATTTC-----TGTCAAGCTT 1118
   |||||
QY 158 LeuAlaAla 160
   |||||
DB 1117 CTTCTGGCG 1109
   |||||

RESULT 37
US-08-023-764B-1/c
; Sequence 1, Application US/08023764B
```

```
Patent No. 5679540
; GENERAL INFORMATION:
; APPLICANT: Rasmussen, Beth Ann
; TITLE OF INVENTION: Cloning and Identification of a Two
; TITLE OF INVENTION: Component Signal Transducing Regulatory System from
; TITLE OF INVENTION: Bacteroides Fragilis
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSER: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/023,764B
; FILING DATE: 26-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 31,658-00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 831-3246
; TELEFAX: (201) 831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2651 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-023-764B-1

Alignment Scores:
Pred. No.: 0.708 Length: 2651
Score: 86.00 Matches: 45
Percent Similarity: 34.53% Conservative: 32
Best Local Similarity: 20.18% Mismatches: 50
Query Match: 8.94% Indels: 96
DB: 1 Gaps: 10

US-10-627-571-2 (1-188) x US-08-023-764B-1 (1-2651)
QY 31 ThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTyraArgValThrArgGlu 50
   |||||
DB 1768 ACTCTCCGCCCGGATGATGCTCCCTTAATGATCTTAATCTTTTTCACATTAAGCCAGTCC 1709
   |||||
QY 51 TyrThrGln-----AsnLysLysGluAlaGluLysLysIleLys 63
   |||||
DB 1708 CAGACCGGAACCTTTTACATCGTCAGATTACCTGTATGACACGCGATGAGAACTTATCAAA 1649
   |||||
QY 64 AsnLysLysLysThrValIle-----LysLeuAlaIleLeuTyraArgAsnAsnGln 80
   |||||
DB 1648 CACCTTCTTCAGGTTTCTTTTAAATACCAATACGTTGCTCTGTATCGAATATCATACG 1589
   |||||
QY 81 Phe-----AsnGlnAspGluLeuAlaLeuMetGlu 90
   |||||
DB 1588 TTTCACGGGTTCCGTTCCAGGTTCTACGTCGACACGAGTCTTCTTCGGGCTTCTT--- 1532
   |||||
QY 91 LysPheLysLysLysValHisGlnLeuAla----- 100
   |||||
DB 1531 GATATTTCACCGGTTATCCATCAGTTGAATATCACTTGTCATATGCAATTTCGTCGCC 1472
   |||||
QY 101 -----MetThrValIleSerPheHis 107
   |||||
```

Db 1471 AATATTAACAGATTGTAGCTCAAGTTCGATGTATCTTACATTATAGGCTTCAC 1412
 Qy 107 ----- 107
 Db 1411 CTTCAGACGACGATTTGATTAACCCGAAATCACTGATGACATCGATTCTTTCAAT 1352
 Qy 108 ----- 120
 Db 1351 CTTCAGTGTCTTCTTCTGTCTGTGACATGACATGTGAAGAACTTTCTCCACCTGGAA 1292
 Qy 121 ----- 132
 Db 1291 TCTCAACGCTGTTCATCATTTAATGACTCCGATATATGCTGGAACATCTCGCGTGA 1232
 Qy 133 ----- 144
 Db 1231 TTTCGGAATGCGGATCTTTCAGACATCTGCGGCAAGAGATGTATATCGGTGT 1172
 Qy 145 ----- 157
 Db 1171 CTTCGAACCTGTGTCTATTTGTTGATTAAGTCACTTCTTCAATTC-----TCTCAGCTT 1118
 Qy 158 LeuAla160
 Db 1117 CTTCGTGCGC 1109

RESULT 38
 US-09-134-000C-1489
 ; Sequence 1489, Application US/09134000C
 ; Patent No. 6617156
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; FILE REFERENCE: 032796-032
 ; CURRENT APPLICATION NUMBER: US/09/134,000C
 ; PRIOR FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/055,778
 ; NUMBER OF SEQ ID NOS: 6812
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1489
 ; LENGTH: 618
 ; TYPE: DNA
 ; ORGANISM: Enterococcus faecalis
 US-09-134-000C-1489

Alignment Scores:
 Pred. No.: 0.0922 Length: 618
 Score: 85.50 Matches: 39
 Percent Similarity: 43.21% Conservative: 31
 Best Local Similarity: 24.07% Mismatches: 52
 Query Match: 8.89% Indels: 41
 DB: 4 Gaps: 6

US-10-627-571-2 (1-188) x US-09-134-000C-1489 (1-618)
 Qy 5 ValPheAsnSerIyAsnLeuAlaValGlnAlaGlnIySerIleLeuGlyLysMetVal 24
 Db 25 ATTATCAACGACGACATTTTGTCTTCAAAATCAAAAACCAATGAAGG----- 75
 Qy 25 SerIySerIleAlaThrThrIleuIleAspAspThrSerSerGluValLeuAspGluLeu 44
 Db 76 -----ATTACGGTTGACTTATGATACATTAAGCGTTGAATTAATTAAGAAAGA 123
 Qy 45 TyrArgValThrArgGlu-----TyrThrGlnAsn 54
 Db 124 TATGAATTTTCAGAGAGAAAAAATCTGGTATTTCTCGAAAAAGTGAATGACGCAAGCG 183
 Qy 55 IyAlaGluAlaGluIySerIleIyAsnLeuIleIyThrValIleIyLeuAlaIle 74
 Db 184 GCACCTTCATCGAAGAAATATCAAAAGAAATTAAGAGCATCAAGCCATT----- 237

Qy 75 LeuTyrArgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGluIySerPheIyLys 94
 Db 238 -----GGAAGAAATTTTACGAAATGATGATTTGAATGATGCTTCTGTAAGAA 288
 Qy 95 IyValHisGlnIleuAlaMetThrValValIserPheHisGlnValAspTyrThrPheAsp 114
 Db 289 GCGGTG-----TATGATAGCATTTAGAGAAACGG--GGA 319
 Qy 115 ArgAsnValLeuSerArgLeuLeuAsnGluCyArgGluMetLeuHisGlnIle----- 132
 Db 320 CGATTGACATGACGAGATTGCTGACCAAGTGTTAAGAAATGATATCAGTAAATTAAG 379
 Qy 133 -----IleGlnArgHisLeuThrAlaIySerHisGlyArgValAsn 146
 Db 380 CTTCACCAAGAAAGTGAATGATGATGAAATTGTTCACAAA--CACCGCCAGTCAGAG 436
 Qy 147 AsnVal 148
 Db 437 AATCTC 442

RESULT 39
 PCT-US94-01101-1/C
 ; Sequence 1, Application PC/TUS9401101
 ; GENERAL INFORMATION:
 ; APPLICANT:
 ; APPLICANT:
 ; APPLICANT:
 ; APPLICANT:
 ; TITLE OF INVENTION: HIGH MOLECULAR WEIGHT B-CELL GROWTH
 ; NUMBER OF SEQUENCES: 8
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
 ; STREET: 600 ATLANTIC AVENUE
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/01101
 ; FILING DATE: FILED HEREWITH
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/005,156
 ; FILING DATE: 15-JAN-1993
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: GATES, EDWARD R.
 ; REGISTRATION NUMBER: 31,616
 ; REFERENCE/DOCKET NUMBER: B0819/7000WO
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/720-3500
 ; TELEFAX: 617/720-2441
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1854 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULAR TYPE: cDNA to mRNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: HOMO SAPIENS
 ; CELL TYPE: LYMPHOMA
 ; PCT-US94-01101-1

Alignment Scores:

Alignment Scores:

Pred. No.:	0.111	Length:	571
Score:	84.50	Matches:	47
Percent Similarity:	41.05%	Conservative:	31
Best Local Similarity:	24.74%	Mismatches:	79
Query Match:	8.78%	Indels:	33
DB:	3	Gaps:	8

US-10-627-571-2 (1-188) X US-09-404-879A-84 (1-571)

```

OY 7 AsnSerIysAsnLeuAlaValGlnAlaGlnIlyLeuGlyIysMetValSerIys 26
Db 505 AACACGAAAGACAAAGAACTTCAGCAACTTCAGCAAACTTGACATCTGATCCAG 444
OY 27 SerIleAlaThrThr-----LeuIleAspAspThrSerSerGluValLeuAsp 42
Db 445 CTTGCACGCTTTACTAAGACATGTCTCCCTCCAGATGATCGTACAGGGTGATAGT 384
OY 43 GluLeuTyArgValThrArgGluTyThrGln-----AsnIlyLeuGluAlaGluIly 60
Db 385 GAACCTACGAAGAAATGGAGAGAAAGTTAGATGATGCATTCAAAGCAAGAAAGAAATT 324
OY 61 LysIleIysAsnLeuIleIysthValIleIyLeuAlaIleLeuTyArgAsnAsnGln 80
Db 325 AGACTCCAAAGACATTAATTGCAGGTGTTTAAAGATCAATCTT-----AGACAGATGCC 274
OY 81 PheAsnGlnAspGluLeuAlaLeu-----MetGluIysPheIysIysVal 96
Db 271 ATCATATGGAAGATTTAAAGATTACATTTCAGGCTTGAAACATGACAAAGCAAGATTGG 214
OY 97 HisGlnLeuAlaMetThrValValSerPheHisGlnValAspTyThrPheAspArgAsn 114
Db 211 GAGTCCCAAGGCCCCAGACAGAGTCTCAGCTTCACAGC-----AAG 174
OY 117 ValLeuSerArgLeuLeuAsnGluIyCysValArgGluMetLeuHisGlnIleLeuGln--Arg 134
Db 172 GTCTGTGATATCTGTACAGGGGAGAAACAAAGAACTTTTGTCCAGGTACAAAGACAGC 114
OY 136 HisLeuThrAlaIySerHisGlyArgValAsnAsnValPheAspIlePheSerAspCys 154
Db 112 CACCTTAATAC-----CACGTTCTCAGAAAT 89
OY 156 GluPheLeuAlaAlaLeuTyArgAsnProPheGluAsnPheIysProHisLeuGlnIlyLeu 174
Db 88 GAA---TTAGCTAAGTTGGAATCAGAACTTAAGAAGTCTCAAGACACAGTTGACTATTTA 324
OY 176 CysAspGlyIleAsnIlyMetLeuAspGlu 185
Db 31 AGTAACCTTTAGAAAATGTAAGCAACA 2

```

RESULT 44

```

US-09-338-933-84/C
; Sequence 84, Application US/09338933
; Patent No. 6488931
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer Lynn
; APPLICANT: King, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: 210121.462C1
; CURRENT APPLICATION NUMBER: US/09/338,933
; CURRENT FILING DATE: 1999-06-23
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 84
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-338-933-84

```

Alignment Scores:

Pred. No.:	0.111	Length:	571
Score:	84.50	Matches:	47

Percent Similarity:	41.05%	Conservative:	3
Best Local Similarity:	24.74%	Mismatches:	7
Query Match:	8.78%	Indels:	3

US-10-627-571-2 (1-188) x US-09-338-933-84 (1-571)

OY		7	AAsenrLyvAenLeuAlaValGlnIlaGlnLysVlySileLeuGlyLysMetValSerLys	26
Dd		505	AAACAAGAAAGACAGAAGAACTTCAGCACTTCCAGAAAACTTGACAGATCTGTGACCCAG	446
OY		27	SerllealathrThr-----LeulleasprThrSerSerluValleAep	42
Dd		445	CTTGAGGCCTTTACTAAGACGATGTCCTCCGCCAGATGATCGTCACAGGGGATGAAT	386
OY		43	GluLeuTyArgValThrArgGluTyTrhGln-----AsnLysGlnAlaGlnLys	60
Dd		385	GAAGCTTAAGAAATGGAGAGAAATTGTATGATGCGATTCAAAAGCAAAGAGAAAT	326
OY		61	LysIleLysAenLeuIleLysThrValIleLysVleAlaIleLeuTyTrhArgAenGln	80
Dd		325	AGACTCCAAGAAGATTAATTCAGTGTCTTAAGATCAACTT----AGACAGATGTCC	272
OY		81	PheaGlnAepGluLeuAlaleu-----MetGluLysPheLysLysVal	96
Dd		271	ATCCCATATGGAAGATTAAGATTAACATTTCCAGGCTTGAACATGACAAAGCAGATTGG	212
OY		97	HisGlnLeuAlaMetThrValLysSerPheHisGlnValAspTyThrPheAspArgAen	116
Dd		211	GAGTCCAAAGCCCCACAGACAGGTCCAGCTTCAGCAG-----AAG	173
OY		117	ValLeuSerAlaGluLeuAenGluCySarArgGluMetLeuHisGlnIleIleGln---Arg	135
Dd		172	GTCGTGTACTCTTACAGGGGAAAACAAAGAACTTTGTGCCACGCTTAAGAGACACCG	113
OY		136	HisLeuThrAlaLysSerHisGluTyArgValAsnAenValPheAspHisPheSerAspCys	155
Dd		112	CACCTATAC-----CACAGTTCTCAGAAT	89
OY		156	GluPheLeuAlaValLeuTyAsnProPheGlyAenPheLysProHisLeuGlnLysLeu	175
Dd		88	GAA---TTAGCTAACTGGAAATCAAGACTTAAGAGCTCAAAAGACAGATTGACTGATTTA	32
OY		176	CysAspGlyIleAsnLysMetLeuAspGlu	185
Dd		31	AGTACTCTTTAGAAAAATGTAAAGAACAA	2

RESULT 45

```

US-09-215-681-84/c
; Sequence 84, Application US/09215681A
; Patent No. 6528253
; GENERAL INFORMATION:
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Fridakis, Tony N.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSTICS
; FILE REFERENCE: 210121.463
; CURRENT APPLICATION NUMBER: US/09/215,681A
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 310
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 84
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-215-681-84

```

Alignment Scores:

Prod. No.:	0.111	Length:	571
Score:	84.50	Matches:	47
Percent Similarity:	41.05%	Conservative:	31
Best Local Similarity:	24.74%	Mismatches:	79

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 17:21:24 ; Search time 502 Seconds

(without alignments)

2216.955 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962

Sequence: 1 MATDVNSKNIAVQAOKIL.....KPHLOKLCGSGINMLDENI 188

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Command line parameters:

-MODE=fixme+ p2n.model -DEV=xlh
-Q/cgnt2.1/USPTO.spool.h/US1067571/tunat.27072005.154719.27497/app.query.fasta_1.327
-DB=N Geneseg -OPMT=fastcap -SUFFIX=ring -MINMATCH=0.1 -LOOFC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1067571 @CGN 1.1 470 @tunat.27072005.154719.27497 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_16Dec04:*
1: genesegm1980s:*
2: genesegm1990s:*
3: genesegm2000s:*
4: genesegm2001as:*
5: genesegm2001bs:*
6: genesegm2002as:*
7: genesegm2002bs:*
8: genesegm2003as:*
9: genesegm2003bs:*
10: genesegm2003cs:*
11: genesegm2003ds:*
12: genesegm2004as:*
13: genesegm2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	954	99.2	1892	8	ACC81105 Human TNF
2	954	99.2	1915	6	ABQ79541 Human SCC
3	954	99.2	1921	4	AAH14888 Human CDN
4	950	98.8	1729	4	AAH14793 Human CDN
5	950	98.8	1814	13	ADR14214 Human NF-

6	950	98.8	2003	13	ADR14216 Human NF-
7	950	98.8	2034	11	ACN44179 Human mRN
8	950	98.8	2081	13	ADR14038 Human NF-
9	950	98.8	58723	11	ACN44178 Human gen
10	948	98.5	1943	2	AAAX8013 Human sec
11	948	98.5	1943	2	ADAI1570 Human CDN
12	948	98.5	1943	10	ADA56241 Gene enco
13	920	95.6	2087	11	ACN44177 Human gen
14	920	95.6	62231	11	ACN44176 Human gen
15	916	95.2	816	4	AAH07403 Human CDN
16	914	95.0	544	12	ACH89719 Human gen
17	651	67.7	1986	12	ADQ63108 Novel hum
18	640	66.5	969	10	ADC30769 Human nov
19	610	63.4	527	12	ACH92216 Human gen
20	593	61.6	619	6	AB878742 DNA encod
21	593	61.6	645	6	AB878741 DNA encod
22	593	61.6	2108	6	AAI39626 Human sec
23	593	61.6	2186	12	ADQ85659 Human tum
24	593	61.6	2186	12	ADQ86751 Human tum
25	593	61.6	2186	13	ADQ84271 Human tum
26	587	61.0	587	4	AAH07641 Human CDN
27	527	54.8	1113	6	ABK34943 Human LP2
28	527	54.8	1165	6	AAO38699 Human CDN
29	527	54.8	1175	6	ABZ11655 Human pol
30	527	54.8	1175	12	ADM44173 Human gen
31	527	54.8	1268	4	AAE59594 Human cel
32	401	41.7	829	4	ABL07183 Drosophi1
33	383.5	39.9	4406	9	ACH48309 Human lun
34	324	33.7	422	9	ACH76016 Human gen
35	304	31.6	500	12	AAH04721 Human CDN
36	292	30.4	714	4	AAH18715 Human CDN
37	292	30.4	1602	4	AAK56622 Human imm
38	281	29.2	441	4	AAK56622 Human sec
39	250	26.0	340	3	AAK10352 Human sec
40	206	21.4	400	4	AAI92435 Human pol
41	166	17.3	252	3	AAK10359 Human sec
42	148	15.4	533	12	ACH78513 Human gen
43	106	11.0	1339	12	ADP04796 Human gen
44	104	10.8	2943	3	AAH01229 Plasmodiu
45	96.5	10.0	588	3	AAH95803 Tomato B-
46	96.5	10.0	1694	3	AAH95805 Tomato B-
47	96.5	10.0	2154	3	AAH95812 Tomato B-
48	95	9.9	638	6	ABK79037 Bacillus
49	93.5	9.7	1809	3	AAH0141 Plasmodiu
50	93.5	9.7	2343	8	ACA22957 Prokaryot
51	93.5	9.7	110000	2	AAK20248_08 Continuation (9 of
52	92.5	9.6	2501	11	ACN92808 Breast ca
53	92	9.6	265118	5	AAH41227 Continuation (7 of
54	91.5	9.5	110000	2	AAK20248_06 Continuation (7 of
55	91.5	9.5	111309	2	AAK20250 Borrelia
56	90.5	9.4	848	4	AAH03594 Human CDN
57	90.5	9.4	1359	8	ACA47387 Prokaryot
58	90.5	9.4	1628	4	AAH13782 Human CDN
59	90.5	9.4	2222	4	AAH75522 Human myo
60	90.5	9.4	2322	2	AAV83134 Polynucle
61	90.5	9.4	2322	6	ABQ92061 Human pol
62	90.5	9.4	4256	5	ADL63542 Human ova
63	90.5	9.4	4428	11	ADH39127 Cancer/an
64	90.5	9.4	4500	10	ADH15555 Human sec
65	90.5	9.4	4501	13	ACN39836 Tumour-as
66	90.5	9.4	5923	12	ADL13178 Human ste
67	90.5	9.4	6633	8	ABX34762 Human mdd
68	90.5	9.4	110000	10	ADP77343 Lactic ac
69	90	9.4	1887	2	AAK99539 Nucleic ac
70	90	9.4	110000	2	AAK99539 Nucleic ac
71	89.5	9.3	1140	2	AAK56021 Merozoite
72	89.5	9.3	1140	2	AAK56021 Merozoite
73	89.5	9.3	1142	10	ABX15269 P. falci
74	89.5	9.3	1142	11	ADM86686 Plasmodiu
75	89	9.3	6538	3	AAK77192 Human ORF
76	88.5	9.2	1351	4	ABL24777 Drosophi1
77	88.5	9.2	1545	4	ABL19371 Drosophi1
78	88.5	9.2	1700	6	ABL63353 Breast ca

79	88.5	9.2	1700	6	ABK64326	Abk64326 Human cDN
80	88.5	9.2	1700	10	AC556483	AC556483 Human sig
81	88.5	9.2	1700	11	AD131572	AD131572 Human cDN
82	88.5	9.2	1700	12	AD156279	AD156279 Human pol
83	88.5	9.2	3204	4	AAH54121	AAH54121 S. epider
84	88.5	9.2	3227	4	AAH54947	AAH54947 S. epider
85	88.5	9.2	3760	4	AAH54665	AAH54665 S. epider
86	88.5	9.2	4302	4	AB124776	AB124776 Drosophila
87	88.5	9.2	4396	4	AB119370	AB119370 Drosophila
88	88	9.1	1638	5	AAH26501	AAH26501 Human low
89	88	9.1	1793	8	AB234807	AB234807 Coding se
90	88	9.1	1993	12	AD078021	AD078021 Human syn
91	88	9.1	2188	4	AAK51890	AAK51890 Human pol
92	88	9.1	2523	4	ABA09008	ABA09008 Human LDL
93	88	9.1	2523	4	AAK52874	AAK52874 Human pol
94	88	9.1	3500	10	ADC39173	ADC39173 Novel hum
95	88	9.1	4206	13	ADR75291	ADR75291 Bovine ac
96	88	9.1	4697	2	AAV32839	AAV32839 Human low
97	88	9.1	4697	5	AAH26496	AAH26496 Human low
98	88	9.1	4730	13	ADR75289	ADR75289 Bovine ac
99	88	9.1	5085	8	ACC46153	ACC46153 Human dlt
100	88	9.1	6577	8	ACC46130	ACC46130 Human dlt

ALIGNMENTS

```
RESULT 1
ACC81105 1 ACC81105 standard; mRNA; 1892 BP.
AC      ACC81105;
XX
XX      25-JUL-2003 (first entry)
DE      Human TNF-induced protein GG2-1 mRNA.
XX
XX      Human: T lymphocyte activation; T-cell; A-raf-1; TCPTP/PTPN2; asthma;
XX      immunosuppressive; antiallergic; antiinflammatory;
XX      lymphocyte activation; lymphocyte migration; cytokine production;
XX      cell surface marker expression; antibody production; apoptosis; allergy;
XX      antibody proliferation; antibody differentiation; hypersensitivity;
XX      graft versus host disease; inflammation; gene; ss; TNF-induced protein;
XX      GG2-1.
XX
XX      Homo sapiens.
OS
XX      WO2003029277-A2.
XX
XX      10-APR-2003.
XX
XX      02-OCT-2002; 2002WO-US031618.
XX
XX      03-OCT-2001; 2001US-0327212P.
XX
XX      (RIGE-) RIGEL PHARM INC.
XX
XX      Chu P, Li C, Liao XC, Masuda E, Pardo J, Zhao H;
XX      WPI; 2003-363276/34.
XX      P-PSDB; ABR59710.
XX
XX      Identifying a compound that modulates T lymphocyte activation, useful for
XX      monitoring changes in cell surface marker expression, comprises
XX      contacting a T cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with
XX      a compound.
XX
XX      Disclosure; Page 84; 126pp; English.
XX
XX      The invention relates to a novel method for identifying a compound that
XX      modulates T lymphocyte activation. The method comprises contacting a T
XX      cell comprising an A-raf-1 or TCPTP/PTPN2 polypeptide with a compound,
XX      where the A-raf-1 or TCPTP/PTPN2 polypeptide is encoded by a nucleic
XX      acid that hybridizes to a nucleic acid encoding a polypeptide having a
```

```
CC      sequence selected from two 606-amino acid sequence and a 415-amino acid
CC      sequence given in the specification. The method of the invention has
CC      immunosuppressive, antiallergic, and antiinflammatory
CC      activity. The method is useful for identifying compounds that modulate
CC      lymphocyte activation and migration, and for monitoring changes in cell
CC      surface marker expression, cytokine production, antibody production,
CC      proliferation and differentiation, and apoptosis, using either cell lines
CC      or primary cells. The A-raf-1 or TCPTP/PTPN2 proteins may be used as
CC      drug targets for compounds that suppress or activate lymphocyte
CC      activation and migration, e.g. for the treatment of diseases in which
CC      modulation of the immune response is desired such as delayed type
CC      hypersensitivity reactions, asthma, allergies, graft versus host disease,
CC      and acute and chronic inflammation. Modulators of lymphocyte activation
CC      are useful for treating disorders related to T and B cell activation and
CC      migration. The present sequence is used in the exemplification of the
CC      invention.
XX
XX      SQ      Sequence 1892 BP; 618 A; 318 C; 381 G; 575 T; 0 U; 0 Other;
XX
XX      Alignment Scores:
XX      Pred. No.:      4,14e-98      Length:      1892
XX      Score:      954.00      Matches:      187
XX      Percent Similarity:      99.47%      Conservative:      0
XX      Best Local Similarity:      99.47%      Mismatches:      1
XX      Query Match:      99.17%      Indels:      0
XX      DB:      8      Gaps:      0
XX
XX      US-10-627-571-2 (1-188) x ACC81105 (1-1892)
XX
XX      QY      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaGlnAlaGlnLysLysLeu 20
XX      DB      98 ATGGCCACGAGTGTCTTATTCGAAACCGCGCGTTCAGGCAAAAGAGATCTTG 157
XX
XX      QY      21 GlyLysMetValSerLysSerLeaLThrThrLeuLLeaAspThrSerSerGluVal 40
XX      DB      158 GGTAAAGTGTCTCCAAATCCATCCGACCACTTAATAGACACAAAGTGTGAGTG 217
XX
XX      QY      41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGlnLys 60
XX      DB      218 CTGGAGAGCTTACGAGTGACCGAGGAGTACACCCAAACAGAGAGGACAGAGAG 277
XX
XX      QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
XX      DB      278 ATCATCAAGAACCTCATCAAGCACTCATCAAGCTGGCATCTTATAGATATATCG 337
XX
XX      QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
XX      DB      338 TTAAATCAAGATGACTGATGATGAGAAATTTAAGAGAAAGATTTCATCAGCTTGC 397
XX
XX      QY      101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
XX      DB      398 ATGACCGTGTGATGTTCCATCAGGTGAGTATTAATACCTTGACCGGAATGTATTATCGAG 457
XX
XX      QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
XX      DB      458 CTGTTAAATGATGACAGAGATGCTGCACCAATCATTCAGCGGCACTCATCGCAAG 517
XX
XX      QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
XX      DB      518 TCACATGACGACGGGTAAATATATCTTTGATCATCTTTTTCAGATTGTGAATTTTGGCTGCC 577
XX
XX      QY      161 LeuTyrAsnProPheGluAsnPheLysProHisGlnIleLysLeuCysAspGlyIleAsn 180
XX      DB      578 TTGTATTAATCTTTTGGGAATTTTAAACCCCATTAACAAACATATGATGATATCAAC 637
XX
XX      QY      181 LysMetLeuAspGluGluAsnIle 188
XX      DB      638 AAAATGTTGATGAAGAGAACATA 661
XX
XX      RESULT 2
XX      ABQ79541
XX      ID      ABQ79541 standard; cDNA; 1915 BP.
```

AC AEO79541;
 XX 25-NOV-2002 (first entry)
 XX
 DE Human SCC-S2 protein encoding cDNA.
 XX
 KM SSC-S2; apoptosis; tumour; cancer; cytostatic; antisense therapy; human;
 XX gene; ss.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH CDS 134..700
 FT /*tag= a
 FT /product= "SCC-S2"
 XX
 PN W0200259337-A1.
 PD 01-AUG-2002.
 XX
 PF 28-JAN-2002; 2002MO-US002212.
 XX
 PR 26-JAN-2001; 2001US-0264062P.
 XX
 PA (GEOU) UNIV GEORGETOWN SCHOOL MEDICINE.
 XX
 PI Kasid UN, Kumar D, Gokhale P, Ahmad I;
 XX
 DR WPI; 2002-657478/70.
 DR P-PSDB; ABB81967.
 XX
 PT New SCC-S2 polypeptides and nucleic acids encoding them, useful as a
 PT target for identifying compounds that modulate cancer progression by
 PT inhibiting apoptosis, as a target for detecting cancers, or for
 PT immunizing animals.
 XX
 PS Claim 2; Fig 1; 70pp; English.
 XX
 CC The invention relates to an anti-apoptotic gene SSC-S2 and encoded
 CC protein. The gene is a positive mediator of tumour growth and metastasis
 CC in certain cancer types. The SSC-S2 protein can be expressed by standard
 CC recombinant methodology. The SSC-S2 polypeptide is useful as a target for
 CC identifying compounds that modulate cancer progression by inhibiting
 CC apoptosis, as a target for detecting cancers where this polypeptide is
 CC overexpressed, e.g. renal and ovarian cancers, and leukemia. The antibody
 CC and antisense oligonucleotide can be used to treat cancer and to inhibit
 CC cancer cell proliferation and/or metastasis. The present sequence
 CC represents a human SSC-S2 protein encoding cDNA
 XX
 SO Sequence 1915 BP; 610 A; 331 C; 392 G; 582 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 4.21e-98 Length: 1915
 Score: 954.00 Matches: 187
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 1
 Query Match: 99.17% Indels: 0
 DB: Gaps: 0
 US-10-627-571-2 (1-188) x ABO79541 (1-1915)
 QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAglLysLysIleLeu 20
 DB 134 ATGGCCACAGATCTCTTAATTCCAAAACCTGCGCTCAGGCACAAAGAGATCTTG 193
 QY 21 GlyLysMetValSerLysSerIleAlaThrThreIleAspAspThrSerSerGluVal 40
 DB 194 GGTAAATGTTGCCAATCATCGCCACCACTTAATAGACACACAAATAGAGG 253
 QY 41 LeuAspGluLeuValThrArgValThrArgGluThrGlnAsnLysLysGluAglLys 60
 DB 254 CTGATGAGCTCTACAGAGTGACCAAGGAGTACACCAAAACAAGAAAGAGGAGGAG 313

QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
 DB 314 ATCATCAAGAACTCTATCAAGACAGATCATAGCTGGCCATCTTATAGAAATATCAG 373
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 DB 374 TTTAATCAAGATGAGCTAGCATTCATGAGACAAATTTAAGAAAGAAAGTTCAGCTTCT 433
 QY 101 MetThrValLysPheHisGlnValAspTyrThrPheAspArgAsnValLleSerArg 120
 DB 434 ATGACCGTGTCTCATCTTCATCATCGTGGATTATACCTTGACCGAATGTGTTATCCAG 493
 QY 121 LeuLeuAsnGluCysArgGluMetLysHisGlnIleGlnArgHisLeuThrAlaLys 140
 DB 494 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTTAGCGCACCTCAGTCCAG 553
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 DB 554 TCACATGAGACGGTAAATATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 613
 QY 161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 DB 614 TTGTATATATCTTTTGGAAATTTTAAACCCACTTCAAAAACTATGTGATGATCAAC 673
 QY 181 LysMetLeuAspGluGluLeuIle 188
 DB 674 AAATGTGATGAGAGAGAACATCA 697
 RESULT 3
 ID AAH14888 standard; cDNA; 1921 BP.
 XX
 AC AAH14888;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:12751.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000BP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 XX Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Iehli S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesising polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 12751; 2537PP + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX
XX
SQ Sequence 1921 BP; 610 A; 334 C; 391 G; 586 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4,23e-98	Length:	1921
Score:	954.00	Matches:	187
Percent Similarity:	99.47%	Conservative:	0
Best Local Similarity:	99.47%	Mismatches:	1
Query Match:	99.17%	Indels:	0
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x AAH14888 (1-1921)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeu 20
DB 140 ATGGCCAGAGATGTCTTTAATCCAAAACTGGCCGCTCAGGCACAAAGAAGATCTTG 199
QY 21 GilylYMetValSerLysSerIleAlaThrLysLeuIleAspAspThrSerGluVal 40
DB 200 GGTAAATAGTGTCCAAATCCATCGCCACACTTAATAGACACCAAGTAGTGAGTGTG 259
QY 41 LeuAspGluLeuTyrrArgValThrArgGluTyrrThrGlnAsnLysGlnAlaGluLys 60
DB 260 CTGGATGAGCTCTACAGAGTGCACAGGAGTACACCCAAACAGAAAGGACGACGAGG 319
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnAngin 80
DB 320 ATCATCAAGAACCTCATCAAGACAGCATCAAGCTGCGCATTTCTTATAGCAATATCAG 379
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 380 TTTAATCAAGATGAGCTTACATTGATGAGAAATTAAAGAAAGTTTCATCAGTTGCT 439
QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
DB 440 ATGACCGGTGTCAATTCATCAGGCGATTAACCTTGGACCGGATGTGTTATCCAGG 499
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 500 CTGTTAAATGAATGCGAGAGATGCTGCACCAATATCTCAGGCCACCTCAGTCCAGG 559
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGlnPheLeuAla 160
DB 560 TCACATGAGCGGGTTAATATGTTGATGATTTTTCAGATGTGAATTTTGGGTGCG 619
QY 161 LeuTyrrAsnProPheGlnLysAsnPheLysProHisGlnGlnLysCysAspGlyIleAsn 180
DB 620 TTGTAAATATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGATGATATCAAC 679
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 680 AAAATGTTGATGAGAGAAACATA 703

RESULT 4
AAH14793
ID AAH14793 standard; cDNA; 1729 BP.

XX
XX AAH14793;
AC 26-JUN-2001 (first entry)
XX
XX
DE Human cDNA sequence SEQ ID NO:12578.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1074617-A2.
XX
XX
PD 07-FEB-2001.
XX
XX
PF 28-JUL-2000; 2000BP-00116126.
XX
XX
PR 29-JUL-1999; 99JP-00248036.
XX
XX 27-AUG-1999; 99JP-00300253.
PR 11-JAN-2000; 2000JP-00118776.
XX
XX 02-MAY-2000; 2000JP-00183767.
PR 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
PI Ota T, Isegaki T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI; 2001-318749/34.
XX
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT length cDNAs defined in the specification, and for the detection and/or
PT diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX
XX
PS Claim 8; SEQ ID NO 12578; 2537bp + Sequence Listing; English.
XX
XX
CC The present invention describes primer sets for synthesizing 5602 full-
CC length cDNAs defined in the specification. Where a primer set comprises:
CC (a) an oligo-dT primer and an oligonucleotide complementary to the
CC complementary strand of a polynucleotide which comprises one of the 5602
CC nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention

XX
XX
SQ Sequence 1729 BP; 563 A; 277 C; 354 G; 535 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.05e-97	Length:	1729
Score:	950.00	Matches:	186
Percent Similarity:	99.47%	Conservative:	1
Best Local Similarity:	98.94%	Mismatches:	1
Query Match:	98.75%	Indels:	0
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x AAH14793 (1-1729)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeu 20

```

Db      101 GTGGCCGACAGATGCTTAATTCGAAAAACCTGGCCGCTCAGGACGACAAAAGATCTTG 160
QY      21 GLYVsmetValserlySer11ealathThleu1leasphapThrSerSerg1Val 40
Db      161 GGTAAATGTGTGTCGAATCCATCGCCACCACTTAATAGACGACAAAGTAGAGGTG 220
QY      41 LeuAspGluLeuValThrArgGluThrGlnAsnLysValAsnGluVal 60
Db      221 CTGGATGAGCTCTACAGAGTACCGAGGAGTACCCCAAAACAGAAAGGAGGACGAGAG 280
QY      61 Lys11leYsAsnLeu1leYsThrVal11leYsLeu1al1leuYrArgAsnGln 80
Db      281 ATCATCAAGAACTCATCAAGACAGATCATCAAGCTGCCATCTTTTATAGAAATACAG 340
QY      81 PheAsnGlnAspGluLeuValAlaMetGluLysPheLysLysValH1SglnLeu1a 100
Db      341 TTTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 400
QY      101 MetThrValValSerPheH1SglnValAspThrThrPheAspArgAsnVal1leuSerArg 120
Db      401 ATGACCGTGTGCTGATTTCCATCAGTGGATTTATACCTTTGACCGGAATGTATTATCCAGG 460
QY      121 LeuLeuAsnGluCysArgGluMetLeuH1Sgln1le1SglnArgH1SleuThra1aLys 140
Db      461 CTGTTAAATGAATGACAGAGAGATGCTGCACCAATCATTTACGCGCCACTCATGCGCAG 520
QY      141 SerH1Sg1YArgValAsnAsnVal1PheAspH1SphSerAspCysGluPheLeu1a1a 160
Db      521 TCACATGACGAGCGGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 580
QY      161 LeuThrAspProPheGluAsnPheLysPheH1SleuGlnLysLeuCysAspGly1leAsn 180
Db      581 TGTATAATATCTTTTGGGAATTTTAAACCCCACTTCAAAACATGATGTATGATATCAAC 640
QY      181 LysMetLeuAspGluGluAsn1le 188
Db      641 AAAATGTTGATGAAGAGAAACATA 664

RESULT 5
ADRI4214
ID      ADRI4214 standard; DNA; 1814 BP.
XX
XX
XX      ADRI4214;
XX
XX      21-OCT-2004 (first entry)
XX
XX
XX      Human NF-kappaB pathway-associated gene SegID215.
XX
XX      NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
XX      antiarthritic; antirheumatic; gastrointestinal-Gen; antiaesthetic;
XX      antiarteriosclerotic; immunomodulator; cerebroprotective; vasotrophic;
XX      immunosuppressive; vulnerrary; gene therapy; immune disorder;
XX      inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
XX      hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
XX      hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
XX      X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
XX      viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
XX      rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
XX      atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
XX      autoimmune disorder; hyper immune activity;
XX      aberrant acute phase response; hypercongenital condition; birth defect;
XX      necrotic lesion; wound; organ transplant rejection;
XX      aberrant signal transduction; proliferating disorder; cancer;
XX      HIV propagation; gene; de; human.
OS      Homo sapiens.
XX
XX      MO2004065577-A2.
XX
XX      05-AUG-2004.
XX
XX

```

```

PF      13 -JAN-2004; 2004MO-US0000798.
XX
XX      14 -JAN-2003; 2003US-0440068P.
PR      12 -MAY-2003; 2003US-0469757P.
XX
XX      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX      Nadler SG, Neubauer MG, Feder JN, Carmen J,
XX      WPI; 2004-562168/54.
DR      P-PSDB; ADRI4215.
XX
XX      New isolated polynucleotides and polypeptides associated with NF-kappaB
PT      pathway, useful for diagnosing, treating, or preventing disorders or
PT      diseases associated with NF-kappaB pathway.
XX
XX      Claim 1; SEQ ID NO 215; 237bp; English.
XX
XX      This invention relates to the novel association of protein sequences (and
CC      the genes which encode them) to the NF-kappaB pathway. The invention may
CC      be useful for the production of compounds with an antiinflammatory,
CC      cytostatic, hepatotropic, virucide, antiaesthetic, antirheumatic,
CC      gastrointestinal-Gen, antiaesthetic, antiarteriosclerotic,
CC      immunomodulator, cerebroprotective, vasotrophic, immunosuppressive or
CC      vulnerrary activity or for gene therapy. The proteins and nucleotides are
CC      useful for diagnosing, preventing, treating, or ameliorating conditions
CC      or diseases associated with the NF-kappaB pathway. The condition is an
CC      immune disorder, an inflammatory disorder, an inflammatory disorder
CC      related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC      hepatic disorders, Hodgkin's lymphoma, haematopoietic tumours, hyper-IgM
CC      syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC      ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC      hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC      survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC      bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC      syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC      immune activity, disorders related to aberrant acute phase responses,
CC      hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC      organ transplant rejection, conditions related to organ transplant
CC      rejection, disorders related to aberrant signal transduction,
CC      proliferating disorders, cancers and HIV propagation in cells infected
CC      with other viruses. The present sequence is that of a human gene which is
CC      subject to the novel association with the NF-kappaB pathway of the
CC      invention. Note: This sequence does not appear in the specification but
CC      was obtained by the indexer from Genbank.
XX
XX      Sequence 1814 BP; 590 A; 297 C; 372 G; 555 T; 0 U; 0 Other;
SQ
XX
XX      Alignment Scores:
XX      Pred. No.: 1,11e-97 Length: 1814
XX      Score: 950.00 Matches: 186
XX      Percent Similarity: 99.47% Conservative: 1
XX      Best Local Similarity: 98.94% Mismatches: 1
XX      Query Match: 98.75% Indels: 0
XX      DB: 13 Gaps: 0

US-10-627-571-2 (1-188) x ADRI4214 (1-1814)
QY      1 MetAlaThAspValPheAsnSerlyAsnLeu1aValGlna1aGlnLys1leu 20
Db      174 GTGGCCGACAGATGCTTAATTCGAAAAACCTGGCCGCTCAGGACGACAAAAGATCTTG 233
QY      21 GLYVsmetValserlySer11ealathThleu1leasphapThrSerSerg1Val 40
Db      234 GGTAAATGTGTGTCGAATCCATCGCCACCACTTAATAGACGACAAAGTAGAGGTG 293
QY      41 LeuAspGluLeuValThrArgGluThrGlnAsnLysValAsnGluVal 60
Db      294 CTGGACGAGCTCTACAGAGTACCGAGGAGTACCCCAAAACAGAAAGGAGGACGAGAG 353
QY      61 Lys11leYsAsnLeu1leYsThrVal11leYsLeu1al1leuYrArgAsnGln 80
Db      354 ATCATCAAGAACTCATCAAGACAGATCATCAAGCTGCCATCTTTTATAGAAATACAG 413

```

OY		81	PheAenGlnAspGIUleuaLaleutMetGulySpHeLylySValhiSGinLeuaLa	100
Db		414	TTTTAAATCAAGAGACGTACCATTGAAGAATAATTMAACAAGAAAGTTTCATCAGCTTGCT	473
OY		101	MettTrAlValSeSPheHisgInVaLaAPYrThPheaSprgaSenVallEuSerArg	120
Db		474	ATGACCGTGGTCAAGTTCCATCAGGTGGATTAATCCTTTGACCGGAAATGTTATCAAG	533
OY		121	LeuleuAnGluCyArGlUwMeIueHISginLieISglNaRghIsLeuthraLaLyS	140
Db		534	CCTGTAATTAATGCAGAGAGATGCTGCACCAAATCATTCAGGCCACCTCACCTGCCAAG	593
OY		141	SerHiISgIVArGVaLasasentValPheaDPhisPhesSetASPcSeglUpheLeuaLaIA	160
Db		594	TCAACATGACGGGTAAATATATGCTTGTATCAATTTTTCAGATTGTAAATTTTGGCTGCC	653
OY		161	LeuTYrasNPProPhEGLYasnPHelysPROHIslEUGLIuLyseuCysaSPGLyilean	180
Db		654	TTGTMTAATCCTTTTGGGAATTTTAAACCCACTTCAAAAATAIGTAGTATCMAC	713
OY		181	LysMetLeuaspGluGUbaSenlle	188
Db		714	AAAAATGTTGATGAAGAGAACAATA	737
RESULT 6				
ID	ADRI4216		standard; DNA; 2003 BP.	
XX	AC	ADR14216;		
XX	DT	21-OCT-2004	(first entry)	
DE			Human NF-kappaB pathway-associated gene SegID217.	
XX			NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;	
KW			antiarthritic; antirheumatic; gastrointestinal-gen; antiaesthetic;	
KW			antiatherosclerotic; immunomodulatory; cerebroprotective; vasotrophic;	
KM			immunosuppressive; vulnerrary; gene therapy; immune disorder;	
KM			inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;	
KW			hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;	
KW			hyper-Igm syndrome; hypohidrotic ectodermal dysplasia;	
KM			X-linked anhidrotic ectodermal dysplasia; immunodeficiency;	
KW			viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;	
KW			viral replication; host cell survival; evasion of immune response;	
KM			rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;	
KW			atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;	
KM			autoimmune disorder; hyper immune activity;	
KW			aberrant acute phase response; hypercongenital condition; birth defect;	
KW			necrotic lesion; wound; organ transplant rejection;	
KW			aberrant signal transduction; proliferating disorder; cancer;	
KX			HIV propagation; gene; ds; human.	
OS			Homo sapiens.	
XX				
PN		WO2004065577-A2.		
XX				
PD		05-AUG-2004.		
XX				
PF		13-JAN-2004; 2004WO-US000798.		
XX				
PR		14-JAN-2003; 2003US-0440068B.		
PR		12-WAY-2003; 2003US-0469757P.		
XX				
PA			(BRIM) BRISTOL-MYERS SQUIBB CO.	
XX				
PI			Nadler SG, Neubauer MG, Feder JN, Carman J,	
DR		WPI; 2004-562168/54.		
PT		P-PDB; ADRI4217.		
FT		New isolated polynucleotides and polypeptides associated with NF-kappaB		
XX		pathway, useful for diagnosing, treating, or preventing disorders or		

PT	diseases associated with NF-kappa pathway.
XX	Claim 1; SEQ ID NO 217; 237bp; English.
PS	This invention relates to the novel association of protein sequences (and
CC	the genes which encode them) to the NF-kappaB pathway. The invention may
CC	be useful for the production of compounds with an antiinflammatory,
CC	cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
CC	gastrointestinal-Gen, antiallergic, antidiabetic,
CC	immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
CC	ulnervary activity or for gene therapy. The proteins and nucleotides are
CC	useful for diagnosing, preventing, treating, or ameliorating conditions
CC	or diseases associated with the NF-kappaB pathway. The condition is an
CC	immune disorder, an inflammatory disorder, an inflammatory disorder
CC	related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC	hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
CC	syndromes, hypohidrotic ectodermal dysplasia, X-linked andidrotic
CC	ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC	hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell,
CC	survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC	bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick,
CC	syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC	immune activity, disorders related to aberrant acute phase responses,
CC	hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC	organ transplant rejection, conditions related to organ transplant
CC	rejection, disorders related to aberrant signal transduction,
CC	proliferating disorders, cancers and HIV propagation in cells infected
CC	with other viruses. The present sequence is that of a human gene which is
CC	subject to the novel association with the NF-kappaB pathway of the
CC	invention. Note: This sequence does not appear in the specification but
CC	was obtained by the Indexer from Genbank.
XX	
SEQ	Sequence 2003 BP; 667 A; 324 C; 406 G; 606 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	1,27e-97 Length: 2003
Score:	950.00 Matches: 186
Percent Similarity:	99.47% Conservative: 1
Best Local Similarity:	98.94% Mismatches: 1
Query Match:	98.75% Indels: 0
DB:	13 Gaps: 0
US-10-627-571-2 (1-188) x ADR14216 (1-2003)	
QY	1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleu 20
DB	204 GTGGCCACAGATGCTTTTAATTCGAAAACTGGCGGCTTCAGGCACAAAGAAATCTTG 263
QY	21 GlyLysMetValSerIysSerIleAlaThrThrIleuIleAspThrSerSerGluVal 40
DB	264 GGTAAATGTGTGCMAATTCATCGGCACACCTTATATAGACACACAAAGTAGAGGTG 322
QY	41 LeuAspGluLeuTyrArgValIthrArgGluTyrThrGlnAsnIysGluAlaGluIys 60
DB	324 CTGATGATGCTTACAGAGTGCACAGGGAGTACACCCAAAACAAGAAAGAGCGAAG 383
QY	61 LysIleIysAsnLeuIleIysThrValIleIysLeuAlaIleLeuTyrArgAsnAsnGln 80
DB	384 ATCATCAAGAACTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATGGAATAATCAG 443
QY	81 PheAsnGlnAspGluLeuAlaLeuMetGluIysPheIysIlyIysValIhIsgIleuAla 100
DB	444 TTTATATAGATAGCTACCATTTGATGAGAAATTTAAGAAAGAAAGTTCATCACTTGCT 503
QY	101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
DB	504 ATGACCGTGCTGAGTTTCATCAAGGAGGATTTATACCTTTGACCGGAATGTGTATCCAG 563
QY	121 LeuLeuAsnGluIyAspArgIleuMetLeuHisGlnIleIleGlnPargHisLeuThrAlaIys 140
DB	564 CTGTAAATGAATGACGAGAGATGCTGCACCAATCATTCACGGCCACCTCATCGCCAG 623
QY	141 SerIleGlyArgValAsnAsnValPheAspHisPheSerArgSerGluPheLeuAla 160


```

Db      624 TCACATGACGGGTATTAATGCTTTGATCATTTTTCAGATTGTGAATTTGGCTGCC 683
Qy      161 LeuTYrAsnProPheGlyAsnPhelYsProHisLeuGlnYsLeuCyAspGlyIleAsn 180
Db      684 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAAC 743
Qy      181 LysMetLeuAspGluGluAsnIle 188
Db      744 AAAATGTTGATGAGGAACATTA 767

RESULT 7
ACN44179
ID      ACN44179 standard; cDNA; 2034 BP.
AC      ACN44179;
DT      18-NOV-2004 (first entry)
XX      Human mRNA sequence hCT1950131.
XX      Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
XX      Homo sapiens.
XX      MO2003073826-A2.
XX      12-SEP-2003.
XX      28-FEB-2003; 2003WO-US006235.
XX      01-MAR-2002; 2002US-00087192.
XX      (SAGR-) SAGRES DISCOVERY.
XX      Morris DW;
XX      WPI; 2003-328604/31.
XX      Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX      comprises a nucleotide sequence.
XX      Claim 1; SEQ ID NO 497; ODP; English.
XX      The present invention relates to novel DNA and protein sequences which
XX      are associated with carcinomas. The sequences are useful for: (i) for
XX      screening drug candidates; (ii) for screening of bioactive agent capable
XX      of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX      a bioactive agent capable of modulating the activity of CAP; (iv) for
XX      evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX      carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX      carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a bloodip;
XX      (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX      determining Carcinoma Associated (CA) gene copy number. In addition, the
XX      CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX      carcinoma including lymphoma. The present sequence is one such CA gene
XX      CC sequence. Note: This patent is an equivalent to basic patent
XX      CC US2002182586A1, for which no sequence data was published
XX      SO Sequence 2034 BP; 639 A; 352 C; 412 G; 631 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,3e-97 Length: 2034
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 11 Gaps: 0

US-10-627-571-2 (1-188) x ACN44179 (1-2034)
Qy      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnYsIleLeu 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

```

Db      248 GTGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTACGGACAAAAGATCTTG 307
Qy      21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSergIVal 40
Db      308 GGTAATAATGGGTGTCAAATTCATCGCCACCACTTAATAGACACAAGTAGTGAGGTG 367
Qy      41 LeuAspGluLeuTYrArgValThrArgGluTYrThrGlnAsnYsIleGluVal 60
Db      368 CTGGATGAGCTTACAGAGTACAGGAGTACACCAAAACAAAGAGGAGGAGAAAG 427
Qy      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTYrArgAsnGln 80
Db      428 ATCATCAAGAACTCTATCAAGACAGTCATCAAGCTGCGCATTTTATAGAAATATCAG 487
Qy      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db      488 TTTAATCAAGATGAGTACGATGATGAGAAATTTAAGAGAAATTCATGCTTGGCT 547
Qy      101 MetThrValAlaSerPheHisGlnValAspTYrThrPheAspArgAsnValLeuSerArg 120
Db      548 ATGACCGTGTCAGTTTCATCAGGTCGATTTATACCTTTGACCGGAATGTGTATCCAG 607
Qy      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db      608 CTGTAAATGATGACGAGAGATGCTGCACCAATATTATGCGCACCTCACTGCGAAG 667
Qy      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGlnPheLeuAlaAla 160
Db      668 TCACATGACGCGGTATTAATGATGTTGATCATTTTTCAGATTGTGAATTTGGCTGCC 727
Qy      161 LeuTYrAsnProPheGlyAsnPhelYsProHisLeuGlnYsLeuCyAspGlyIleAsn 180
Db      728 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAAC 787
Qy      181 LysMetLeuAspGluGluAsnIle 188
Db      788 AAAATGTTGATGAGGAACATTA 811

RESULT 8
ADRL4038
ID      ADRL4038 standard; DNA; 2081 BP.
AC      ADRL4038;
DT      21-OCT-2004 (first entry)
XX      Human NF-kappaB pathway-associated gene SegID39.
XX      NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
XX      antiallergic; antirheumatic; gastrointestinal-gen; antiaesthetic;
XX      antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
XX      immunosuppressive; vulnerary; gene therapy; immune disorder;
XX      inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
XX      hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
XX      hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
XX      X-linked ahidrotic ectodermal dysplasia; immunodeficiency;
XX      viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
XX      viral replication; host cell survival; evasion of immune response;
XX      rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
XX      atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
XX      autoimmune disorder; hyper immune activity;
XX      aberrant acute phase response; hypercongenital condition; birth defect;
XX      necrotic lesion; wound; organ transplant rejection;
XX      aberrant signal transduction; proliferating disorder; cancer;
XX      HIV propagation; gene; ds; human.
XX      Homo sapiens.
XX      MO2004065577-A2.
XX      05-AUG-2004.
XX      13-JAN-2004; 2004WO-US000798.

```

XX 14-JAN-2003; 2003US-0440068P.
PR 12-MAY-2003; 2003US-0469757P.
XX
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
PI Nadler SG, Neubauer MG, Feder JN, Carman J,
XX MPI; 2004-562168/54.
DR P-PSDB; ADR14039.
XX
XX New isolated polynucleotides and polypeptides associated with NF-kappa
PT pathway, useful for diagnosing, treating, or preventing disorders or
PT diseases associated with NF-kappaB pathway.
XX
PS Claim 1; SEQ ID NO 39; 237pp; English.

XX This invention relates to the novel association of protein sequences (and
CC the genes which encode them) to the NF-kappaB pathway. The invention may
CC be useful for the production of compounds with an anti-inflammatory,
CC cytostatic, hepatotropic, virostatic, antiarthritic, antirheumatic,
CC gastrointestinal-gen, antiaesthetic, antiarteriosclerotic,
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
CC vulnerrary activity or for gene therapy. The proteins and nucleotides are
CC useful for diagnosing, preventing, treating, or ameliorating conditions
CC or diseases associated with the NF-kappaB pathway. The condition is an
CC immune disorder, an inflammatory disorder, an inflammatory disorder
CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-igm
CC syndromes, hypodiprotic ectodermal dysplasia, X-linked andidrotic
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC immune activity, disorders related to aberrant acute phase responses,
CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, conditions related to organ transplant
CC rejection, disorders related to aberrant signal transduction,
CC proliferating disorders, cancers and HIV propagation in cells infected
CC with other viruses. The present sequence is that of a human gene which is
CC subject to the novel association with the NF-kappaB pathway of the
CC invention. Note: This sequence does not appear in the specification but
CC was obtained by the indexer from Genbank.

XX SQ Sequence 2081 BP; 668 A; 359 C; 422 G; 632 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.34e-97 Length: 2081
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 13 Gaps: 0

US-10-627-571-2 (1-188) x ADR14038 (1-2081)

QY 1 MetAlaThAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlystIleu 20
DB 267 GGGGCCACGAGTCTTAAATTCAAAACCTGGCCGTCAGGCACAAAAGAGATCTTG 326
QY 21 GILYsMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 327 GGTAAATGGTGTCCAAATCCATCCGACACCACTTAATAGACACCAAGTACTGAGTG 386
QY 41 LeuAspGluLeuIlyrArgValTharAgIuYrThGlnAsnIlyysGluAlaGluIys 60
DB 387 CTGGATGAGACTCTACAGAGTGAACAGGGAGTACACCAAAACAAGAGAGCAGGAAG 446
QY 61 LysIleIysAsnLeuIleIysThrValIlyleYsLeuAlaIleuIlyrArgAsnAngin 80
DB 447 ATCATCAAGAACTCTATCAAGACAGTCATCAAGCTGCGCCATCTTATAGAAATATTCAG 506

QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluIysPheIlysIlyValHisGlnLeuAla 100
DB 507 TTTAATCAAGATGACCTAGCATGTGAGAAATTTAAGAAAGATTTCATCAGCTTGT 566
QY 101 MetThrValValSerPheHisGlnValAspIlyrThrPheAspArgAsnValLeuSerArg 120
DB 567 ATGACCGTGGTCCAGTTTCCATCAGGTGATATATACCTTTGACCCGGAATGTGTTATCCAG 626
QY 121 LeuLeuAsnGluIysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrIlyIys 140
DB 627 CTGTTAAATGAATGACAGAGATGCTGCACAAATCATTCACCGCCACTCATCTCCAG 686
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 687 TCACATGACCGCGGTTAATATATGTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 746
QY 161 LeuIlyrAsnProPheGlyAsnPheIysProHisLeuGlnIlyLeuIysAspGlyIleAsn 180
DB 747 TTGTATTAATCCTTTGGGAATTTTAAACCCACTTACAAAACATATGTGATGTATCAAC 806
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 807 AAAATGTTGATGAAGAGACATA 830

RESULT 9
ACN44178
ID ACN44178 standard; DNA; 58723 BP.

ACN44178;

18-NOV-2004 (first entry)

Human genomic sequence hCG36837.

Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

Homo sapiens.

MO2003073826-A2.

12-SEP-2003.

28-FEB-2003; 2003WO-US006235.

01-MAR-2002; 2002US-00087192.

(SAGR-) SAGRES DISCOVERY.

Morris DW;

WPI; 2003-328604/31.

Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
comprises a nucleotide sequence.

Claim 1; SEQ ID NO 496; opp; English.

XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
US2002182586A1, for which no sequence data was published

XX SQ Sequence 58723 BP; 16891 A; 11477 C; 12539 G; 17816 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.08e-95 Length: 58723
 Score: 950.00 Matches: 186
 Percent Similarity: 99.47% Conservative: 1
 Best Local Similarity: 98.94% Mismatches: 1
 Query Match: 98.75% Indels: 0
 DB: 11 Gaps: 0

US-10-627-571-2 (1-188) x ACN44178 (1-58723)

```

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeu 20
   :::::
Db 46937 GTGGCCACAGATGCTTTAATTCGAAAACCTGGCCCTTACGACAAAAGAGATCTTG 46996
QY 21 GlyLysMetValSerLysSerLysLeuAlaThrLeuLeuAspThrSerSerGluVal 40
   :::::
Db 46997 GGTAAATGCTGTCCAAATCCATCGCCACCTTAATAGCAGACAAAGTAGAGGTG 47056
QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLys 60
   :::::
Db 47057 CTGATAGCTCTACAGAGTGAACGAGAGTACACCCAAAACAAGAGGAGGAGAG 47116
QY 61 LysIleLysAsnLeuLysLysThrValLysLeuAlaLysLeuTyrArgAsnAsnGln 80
   :::::
Db 47117 ATCATCAAGAACTCATCAAGACAGCATCATCAAGCTGGCCATCTTTATAGGAATATCAG 47176
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValLysGlnLeuAla 100
   :::::
Db 47177 TTTAATCAAGATGAGTCAAGATTCATGAGAAATTTAAGAAAGTTCATCAGCTTGCT 47236
QY 101 MetThrValValSerPheHISGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
   :::::
Db 47237 ATGACCTGTGCTCAAGTTTCCATCAGTGGATTATACCTTTACCGGAATGTTATCCAG 47296
QY 121 LeuLeuAsnGluCysArgGluMetLeuHISGlnLysGlnArgHISLeuThrAlaLys 140
   :::::
Db 47297 CTGTAAATATAGAGAGAGAGATGCTGACCAATCTTACAGCCCACTCAGCTGCAG 47356
QY 141 SerHISGlyArgValAsnAsnValPheAspHISPheSerAspCysGluPheLeuAlaLys 160
   :::::
Db 47357 TCACATGACGAGGTTAATATGTTGATCATTTTCAATTTGTGAATTTTGGCTGCC 47416
QY 161 LeuTyrAsnProPheGluAsnPheLysPheHISLeuGlnLysLeuCysAspGlyLysAsn 180
   :::::
Db 47417 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAAC 47476
QY 181 LysMetLeuAspGluGluAsnLys 188
   :::::
Db 47477 AAAATGTTGATGAGAGAAACATA 47500

```

RESULT 10

AAK98013

ID AAK98013 standard; DNA; 1943 BP.

XX AAK98013;

XX AC AAK98013;

XX DT 17-SEP-1999 (first entry)

XX DE Human secreted protein gene 98.

XX Human; secreted protein; cancer; tumour; developmental abnormality;
 XX foetal deficiency; blood disorder; immune system disorder; inflammation;
 XX autoimmune disease; allergy; Alzheimer's disease; cognitive disorder;
 XX schizophrania; arthritis; asthma; psoriasis; sepsis; skin disorder;
 XX atherosclerosis; diabetes; cardiovascular disorder; kidney disorder;
 XX digestive disorder; endocrine disorder; infection; AIDS; se.

OS Homo sapiens.

XX WO9931117-A1.

XX 24-JUN-1999.

```

XX PF 17-DEC-1998; 98MO-US027059.
XX PR 18-DEC-1997; 97US-0068006P.
XX PR 18-DEC-1997; 97US-0068007P.
XX PR 18-DEC-1997; 97US-0068008P.
XX PR 18-DEC-1997; 97US-0068009P.
XX PR 18-DEC-1997; 97US-0068010P.
XX PR 18-DEC-1997; 97US-0068011P.
XX PR 18-DEC-1997; 97US-0068012P.
XX PR 18-DEC-1997; 97US-0068013P.
XX PR 18-DEC-1997; 97US-0068014P.
XX PR 18-DEC-1997; 97US-0068015P.
XX PR 18-DEC-1997; 97US-0068016P.
XX PR 18-DEC-1997; 97US-0068017P.
XX PR 18-DEC-1997; 97US-0068018P.
XX PR 18-DEC-1997; 97US-0068019P.
XX PR 18-DEC-1997; 97US-0068020P.
XX PR 18-DEC-1997; 97US-0068021P.
XX PR 18-DEC-1997; 97US-0068022P.
XX PR 18-DEC-1997; 97US-0068023P.
XX PR 18-DEC-1997; 97US-0068024P.
XX PR 18-DEC-1997; 97US-0068025P.
XX PR 18-DEC-1997; 97US-0068026P.
XX PR 18-DEC-1997; 97US-0068027P.
XX PR 18-DEC-1997; 97US-0068028P.
XX PR 18-DEC-1997; 97US-0068029P.
XX PR 18-DEC-1997; 97US-0068030P.
XX PR 18-DEC-1997; 97US-0068031P.
XX PR 18-DEC-1997; 97US-0068032P.
XX PR 18-DEC-1997; 97US-0068033P.
XX PR 18-DEC-1997; 97US-0068034P.
XX PR 18-DEC-1997; 97US-0068035P.
XX PR 18-DEC-1997; 97US-0068036P.
XX PR 18-DEC-1997; 97US-0068037P.
XX PR 18-DEC-1997; 97US-0068038P.
XX PR 18-DEC-1997; 97US-0068039P.
XX PR 18-DEC-1997; 97US-0068040P.
XX PR 18-DEC-1997; 97US-0068041P.
XX PR 18-DEC-1997; 97US-0068042P.
XX PR 18-DEC-1997; 97US-0068043P.
XX PR 18-DEC-1997; 97US-0068044P.
XX PR 18-DEC-1997; 97US-0068045P.
XX PR 18-DEC-1997; 97US-0068046P.
XX PR 18-DEC-1997; 97US-0068047P.
XX PR 18-DEC-1997; 97US-0068048P.
XX PR 18-DEC-1997; 97US-0068049P.
XX PR 18-DEC-1997; 97US-0068050P.
XX PR 18-DEC-1997; 97US-0068051P.
XX PR 18-DEC-1997; 97US-0068052P.
XX PR 18-DEC-1997; 97US-0068053P.
XX PR 18-DEC-1997; 97US-0068054P.
XX PR 18-DEC-1997; 97US-0068055P.
XX PR 18-DEC-1997; 97US-0068056P.
XX PR 18-DEC-1997; 97US-0068057P.
XX PR 18-DEC-1997; 97US-0068058P.
XX PR 18-DEC-1997; 97US-0068059P.
XX PR 18-DEC-1997; 97US-0068060P.
XX PR 18-DEC-1997; 97US-0068061P.
XX PR 18-DEC-1997; 97US-0068062P.
XX PR 18-DEC-1997; 97US-0068063P.
XX PR 18-DEC-1997; 97US-0068064P.
XX PR 18-DEC-1997; 97US-0068065P.
XX PR 18-DEC-1997; 97US-0068066P.
XX PR 18-DEC-1997; 97US-0068067P.
XX PR 18-DEC-1997; 97US-0068068P.
XX PR 18-DEC-1997; 97US-0068069P.
XX PR 18-DEC-1997; 97US-0068070P.
XX PR 18-DEC-1997; 97US-0068071P.
XX PR 18-DEC-1997; 97US-0068072P.
XX PR 18-DEC-1997; 97US-0068073P.
XX PR 18-DEC-1997; 97US-0068074P.
XX PR 18-DEC-1997; 97US-0068075P.
XX PR 18-DEC-1997; 97US-0068076P.
XX PR 18-DEC-1997; 97US-0068077P.
XX PR 18-DEC-1997; 97US-0068078P.
XX PR 18-DEC-1997; 97US-0068079P.
XX PR 18-DEC-1997; 97US-0068080P.
XX PR 18-DEC-1997; 97US-0068081P.
XX PR 18-DEC-1997; 97US-0068082P.
XX PR 18-DEC-1997; 97US-0068083P.
XX PR 18-DEC-1997; 97US-0068084P.
XX PR 18-DEC-1997; 97US-0068085P.
XX PR 18-DEC-1997; 97US-0068086P.
XX PR 18-DEC-1997; 97US-0068087P.
XX PR 18-DEC-1997; 97US-0068088P.
XX PR 18-DEC-1997; 97US-0068089P.
XX PR 18-DEC-1997; 97US-0068090P.
XX PR 18-DEC-1997; 97US-0068091P.
XX PR 18-DEC-1997; 97US-0068092P.
XX PR 18-DEC-1997; 97US-0068093P.
XX PR 18-DEC-1997; 97US-0068094P.
XX PR 18-DEC-1997; 97US-0068095P.
XX PR 18-DEC-1997; 97US-0068096P.
XX PR 18-DEC-1997; 97US-0068097P.
XX PR 18-DEC-1997; 97US-0068098P.
XX PR 18-DEC-1997; 97US-0068099P.
XX PR 18-DEC-1997; 97US-0068100P.
XX PR 18-DEC-1997; 97US-0068101P.
XX PR 18-DEC-1997; 97US-0068102P.
XX PR 18-DEC-1997; 97US-0068103P.
XX PR 18-DEC-1997; 97US-0068104P.
XX PR 18-DEC-1997; 97US-0068105P.
XX PR 18-DEC-1997; 97US-0068106P.
XX PR 18-DEC-1997; 97US-0068107P.
XX PR 18-DEC-1997; 97US-0068108P.
XX PR 18-DEC-1997; 97US-0068109P.
XX PR 18-DEC-1997; 97US-0068110P.
XX PR 18-DEC-1997; 97US-0068111P.
XX PR 18-DEC-1997; 97US-0068112P.
XX PR 18-DEC-1997; 97US-0068113P.
XX PR 18-DEC-1997; 97US-0068114P.
XX PR 18-DEC-1997; 97US-0068115P.
XX PR 18-DEC-1997; 97US-0068116P.
XX PR 18-DEC-1997; 97US-0068117P.
XX PR 18-DEC-1997; 97US-0068118P.
XX PR 18-DEC-1997; 97US-0068119P.
XX PR 18-DEC-1997; 97US-0068120P.
XX PR 18-DEC-1997; 97US-0068121P.
XX PR 18-DEC-1997; 97US-0068122P.
XX PR 18-DEC-1997; 97US-0068123P.
XX PR 18-DEC-1997; 97US-0068124P.
XX PR 18-DEC-1997; 97US-0068125P.
XX PR 18-DEC-1997; 97US-0068126P.
XX PR 18-DEC-1997; 97US-0068127P.
XX PR 18-DEC-1997; 97US-0068128P.
XX PR 18-DEC-1997; 97US-0068129P.
XX PR 18-DEC-1997; 97US-0068130P.
XX PR 18-DEC-1997; 97US-0068131P.
XX PR 18-DEC-1997; 97US-0068132P.
XX PR 18-DEC-1997; 97US-0068133P.
XX PR 18-DEC-1997; 97US-0068134P.
XX PR 18-DEC-1997; 97US-0068135P.
XX PR 18-DEC-1997; 97US-0068136P.
XX PR 18-DEC-1997; 97US-0068137P.
XX PR 18-DEC-1997; 97US-0068138P.
XX PR 18-DEC-1997; 97US-0068139P.
XX PR 18-DEC-1997; 97US-0068140P.
XX PR 18-DEC-1997; 97US-0068141P.
XX PR 18-DEC-1997; 97US-0068142P.
XX PR 18-DEC-1997; 97US-0068143P.
XX PR 18-DEC-1997; 97US-0068144P.
XX PR 18-DEC-1997; 97US-0068145P.
XX PR 18-DEC-1997; 97US-0068146P.
XX PR 18-DEC-1997; 97US-0068147P.
XX PR 18-DEC-1997; 97US-0068148P.
XX PR 18-DEC-1997; 97US-0068149P.
XX PR 18-DEC-1997; 97US-0068150P.
XX PR 18-DEC-1997; 97US-0068151P.
XX PR 18-DEC-1997; 97US-0068152P.
XX PR 18-DEC-1997; 97US-0068153P.
XX PR 18-DEC-1997; 97US-0068154P.
XX PR 18-DEC-1997; 97US-0068155P.
XX PR 18-DEC-1997; 97US-0068156P.
XX PR 18-DEC-1997; 97US-0068157P.
XX PR 18-DEC-1997; 97US-0068158P.
XX PR 18-DEC-1997; 97US-0068159P.
XX PR 18-DEC-1997; 97US-0068160P.
XX PR 18-DEC-1997; 97US-0068161P.
XX PR 18-DEC-1997; 97US-0068162P.
XX PR 18-DEC-1997; 97US-0068163P.
XX PR 18-DEC-1997; 97US-0068164P.
XX PR 18-DEC-1997; 97US-0068165P.
XX PR 18-DEC-1997; 97US-0068166P.
XX PR 18-DEC-1997; 97US-0068167P.
XX PR 18-DEC-1997; 97US-0068168P.
XX PR 18-DEC-1997; 97US-0068169P.
XX PR 18-DEC-1997; 97US-0068170P.
XX PR 18-DEC-1997; 97US-0068171P.
XX PR 18-DEC-1997; 97US-0068172P.
XX PR 18-DEC-1997; 97US-0068173P.
XX PR 18-DEC-1997; 97US-0068174P.
XX PR 18-DEC-1997; 97US-0068175P.
XX PR 18-DEC-1997; 97US-0068176P.
XX PR 18-DEC-1997; 97US-0068177P.
XX PR 18-DEC-1997; 97US-0068178P.
XX PR 18-DEC-1997; 97US-0068179P.
XX PR 18-DEC-1997; 97US-0068180P.
XX PR 18-DEC-1997; 97US-0068181P.
XX PR 18-DEC-1997; 97US-0068182P.
XX PR 18-DEC-1997; 97US-0068183P.
XX PR 18-DEC-1997; 97US-0068184P.
XX PR 18-DEC-1997; 97US-0068185P.
XX PR 18-DEC-1997; 97US-0068186P.
XX PR 18-DEC-1997; 97US-0068187P.
XX PR 18-DEC-1997; 97US-0068188P.
XX PR 18-DEC-1997; 97US-0068189P.
XX PR 18-DEC-1997; 97US-0068190P.
XX PR 18-DEC-1997; 97US-0068191P.
XX PR 18-DEC-1997; 97US-0068192P.
XX PR 18-DEC-1997; 97US-0068193P.
XX PR 18-DEC-1997; 97US-0068194P.
XX PR 18-DEC-1997; 97US-0068195P.
XX PR 18-DEC-1997; 97US-0068196P.
XX PR 18-DEC-1997; 97US-0068197P.
XX PR 18-DEC-1997; 97US-0068198P.
XX PR 18-DEC-1997; 97US-0068199P.
XX PR 18-DEC-1997; 97US-0068200P.
XX PR 18-DEC-1997; 97US-0068201P.
XX PR 18-DEC-1997; 97US-0068202P.
XX PR 18-DEC-1997; 97US-0068203P.
XX PR 18-DEC-1997; 97US-0068204P.
XX PR 18-DEC-1997; 97US-0068205P.
XX PR 18-DEC-1997; 97US-0068206P.
XX PR 18-DEC-1997; 97US-0068207P.
XX PR 18-DEC-1997; 97US-0068208P.
XX PR 18-DEC-1997; 97US-0068209P.
XX PR 18-DEC-1997; 97US-0068210P.
XX PR 18-DEC-1997; 97US-0068211P.
XX PR 18-DEC-1997; 97US-0068212P.
XX PR 18-DEC-1997; 97US-0068213P.
XX PR 18-DEC-1997; 97US-0068214P.
XX PR 18-DEC-1997; 97US-0068215P.
XX PR 18-DEC-1997; 97US-0068216P.
XX PR 18-DEC-1997; 97US-0068217P.
XX PR 18-DEC-1997; 97US-0068218P.
XX PR 18-DEC-1997; 97US-0068219P.
XX PR 18-DEC-1997; 97US-0068220P.
XX PR 18-DEC-1997; 97US-0068221P.
XX PR 18-DEC-1997; 97US-0068222P.
XX PR 18-DEC-1997; 97US-0068223P.
XX PR 18-DEC-1997; 97US-0068224P.
XX PR 18-DEC-1997; 97US-0068225P.
XX PR 18-DEC-1997; 97US-0068226P.
XX PR 18-DEC-1997; 97US-0068227P.
XX PR 18-DEC-1997; 97US-0068228P.
XX PR 18-DEC-1997; 97US-0068229P.
XX PR 18-DEC-1997; 97US-0068230P.
XX PR 18-DEC-1997; 97US-0068231P.
XX PR 18-DEC-1997; 97US-0068232P.
XX PR 18-DEC-1997; 97US-0068233P.
XX PR 18-DEC-1997; 97US-0068234P.
XX PR 18-DEC-1997; 97US-0068235P.
XX PR 18-DEC-1997; 97US-0068236P.
XX PR 18-DEC-1997; 97US-0068237P.
XX PR 18-DEC-1997; 97US-0068238P.
XX PR 18-DEC-1997; 97US-0068239P.
XX PR 18-DEC-1997; 97US-0068240P.
XX PR 18-DEC-1997; 97US-0068241P.
XX PR 18-DEC-1997; 97US-0068242P.
XX PR 18-DEC-1997; 97US-0068243P.
XX PR 18-DEC-1997; 97US-0068244P.
XX PR 18-DEC-1997; 97US-0068245P.
XX PR 18-DEC-1997; 97US-0068246P.
XX PR 18-DEC-1997; 97US-0068247P.
XX PR 18-DEC-1997; 97US-0068248P.
XX PR 18-DEC-1997; 97US-0068249P.
XX PR 18-DEC-1997; 97US-0068250P.
XX PR 18-DEC-1997; 97US-0068251P.
XX PR 18-DEC-1997; 97US-0068252P.
XX PR 18-DEC-1997; 97US-0068253P.
XX PR 18-DEC-1997; 97US-0068254P.
XX PR 18-DEC-1997; 97US-0068255P.
XX PR 18-DEC-1997; 97US-0068256P.
XX PR 18-DEC-1997; 97US-0068257P.
XX PR 18-DEC-1997; 97US-0068258P.
XX PR 18-DEC-1997; 97US-0068259P.
XX PR 18-DEC-1997; 97US-0068260P.
XX PR 18-DEC-1997; 97US-0068261P.
XX PR 18-DEC-1997; 97US-0068262P.
XX PR 18-DEC-1997; 97US-0068263P.
XX PR 18-DEC-1997; 97US-0068264P.
XX PR 18-DEC-1997; 97US-0068265P.
XX PR 18-DEC-1997; 97US-0068266P.
XX PR 18-DEC-1997; 97US-0068267P.
XX PR 18-DEC-1997; 97US-0068268P.
XX PR 18-DEC-1997; 97US-0068269P.
XX PR 18-DEC-1997; 97US-0068270P.
XX PR 18-DEC-1997; 97US-0068271P.
XX PR 18-DEC-1997; 97US-0068272P.
XX PR 18-DEC-1997; 97US-0068273P.
XX PR 18-DEC-1997; 97US-0068274P.
XX PR 18-DEC-1997; 97US-0068275P.
XX PR 18-DEC-1997; 97US-0068276P.
XX PR 18-DEC-1997; 97US-0068277P.
XX PR 18-DEC-1997; 97US-0068278P.
XX PR 18-DEC-1997; 97US-0068279P.
XX PR 18-DEC-1997; 97US-0068280P.
XX PR 18-DEC-1997; 97US-0068281P.
XX PR 18-DEC-1997; 97US-0068282P.
XX PR 18-DEC-1997; 97US-0068283P.
XX PR 18-DEC-1997; 97US-0068284P.
XX PR 18-DEC-1997; 97US-0068285P.
XX PR 18-DEC-1997; 97US-0068286P.
XX PR 18-DEC-1997; 97US-0068287P.
XX PR 18-DEC-1997; 97US-0068288P.
XX PR 18-DEC-1997; 97US-0068289P.
XX PR 18-DEC-1997; 97US-0068290P.
XX PR 18-DEC-1997; 97US-0068291P.
XX PR 18-DEC-1997; 97US-0068292P.
XX PR 18-DEC-1997; 97US-0068293P.
XX PR 18-DEC-1997; 97US-0068294P.
XX PR 18-DEC-1997; 97US-0068295P.
XX PR 18-DEC-1997; 97US-0068296P.
XX PR 18-DEC-1997; 97US-0068297P.
XX PR 18-DEC-1997; 97US-0068298P.
XX PR 18-DEC-1997; 97US-0068299P.
XX PR 18-DEC-1997; 97US-0068300P.
XX PR 18-DEC-1997; 97US-0068301P.
XX PR 18-DEC-1997; 97US-0068302P.
XX PR 18-DEC-1997; 97US-0068303P.
XX PR 18-DEC-1997; 97US-0068304P.
XX PR 18-DEC-1997; 97US-0068305P.
XX PR 18-DEC-1997; 97US-0068306P.
XX PR 18-DEC-1997; 97US-0068307P.
XX PR 18-DEC-1997; 97US-0068308P.
XX PR 18-DEC-1997; 97US-0068309P.
XX PR 18-DEC-1997; 97US-0068310P.
XX PR 18-DEC-1997; 97US-0068311P.
XX PR 18-DEC-1997; 97US-0068312P.
XX PR 18-DEC-1997; 97US-0068313P.
XX PR 18-DEC-1997; 97US-0068314P.
XX PR 18-DEC-1997; 97US-0068315P.
XX PR 18-DEC-1997; 97US-0068316P.
XX PR 18-DEC-1997; 97US-0068317P.
XX PR 18-DEC-1997; 97US-0068318P.
XX PR 18-DEC-1997; 97US-0068319P.
XX PR 18-DEC-1997; 97US-0068320P.
XX PR 18-DEC-1997; 97US-0068321P.
XX PR 18-DEC-1997; 97US-0068322P.
XX PR 18-DEC-1997; 97US-0068323P.
XX PR 18-DEC-1997; 97US-0068324P.
XX PR 18-DEC-1997; 97US-0068325P.
XX PR 18-DEC-1997; 97US-0068326P.
XX PR 18-DEC-1997; 97US-0068327P.
XX PR 18-DEC-1997; 97US-0068328P.
XX PR 18-DEC-1997; 97US-0068329P.
XX PR 18-DEC-1997; 97US-0068330P.
XX PR 18-DEC-1997; 97US-0068331P.
XX PR 18-DEC-1997; 97US-0068332P.
XX PR 18-DEC-1997; 97US-0068333P.
XX PR 18-DEC-1997; 97US-0068334P.
XX PR 18-DEC-1997; 97US-0068335P.
XX PR 18-DEC-1997; 97US-0068336P.
XX PR 18-DEC-1997; 97US-0068337P.
XX PR 18-DEC-1997; 97US-0068338P.
XX PR 18-DEC-1997; 97US-0068339P.
XX PR 18-DEC-1997; 97US-0068340P.
XX PR 18-DEC-1997; 97US-0068341P.
XX PR 18-DEC-1997; 97US-0068342P.
XX PR 18-DEC-1997; 97US-0068343P.
XX PR 18-DEC-1997; 97US-0068344P.
XX PR 18-DEC-1997; 97US-0068345P.
XX PR 18-DEC-1997; 97US-0068346P.
XX PR 18-DEC-1997; 97US-0068347P.
XX PR 18-DEC-1997; 97US-0068348P.
XX PR 18-DEC-1997; 97US-0068349P.
XX PR 18-DEC-1997; 97US-0068350P.
XX PR 18-DEC-1997; 97US-0068351P.
XX PR 18-DEC-1997; 97US-0068352P.
XX PR 18-DEC-1997; 97US-0068353P.
XX PR 18-DEC-1997; 97US-0068354P.
XX PR 18-DEC-1997; 97US-0068355P.
XX PR 18-DEC-1997; 97US-0068356P.
XX PR 18-DEC-1997; 97US-0068357P.
XX PR 18-DEC-1997; 97US-0068358P.
XX PR 18-DEC-1997; 97US-0068359P.
XX PR 18-DEC-1997; 97US-0068360P.
XX PR 18-DEC-1997; 97US-0068361P.
XX PR 18-DEC-1997; 97US-0068362P.
XX PR 18-DEC-1997; 97US-0068363P.
XX PR 18-DEC-1997; 97US-0068364P.
XX PR 18-DEC-1997; 97US-0068365P.
XX PR 18-DEC-1997; 97US-0068366P.
XX PR 18-DEC-1997; 97US-0068367P.
XX PR 18-DEC-1997; 97US-0068368P.
XX PR 18-DEC-1997; 97US-0068369P.
XX PR 18-DEC-1997; 97US-0068370P.
XX PR 18-DEC-1997; 97US-0068371P.
XX PR 18-DEC-1997; 97US-0068372P.
XX PR 18-DEC-1997; 97US-0068373P.
XX PR 18-DEC-1997; 97US-0068374P.
XX PR 18-DEC-1997; 97US-0068375P.
XX PR 18-DEC-1997; 97US-0068376P.
XX PR 18-DEC-1997; 97US-0068377P.
XX PR 18-DEC-1997; 97US-0068378P.
XX PR 18-DEC-1997; 97US-0068379P.
XX PR 18-DEC-1997; 97US-00683
```

```

QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnGln 80
   |||
DB 300 ATCATCAAGAACCTTCAAGACAGTCATCAAGCTTCCATCTTTATAGATATATCCG 359
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAla 100
   |||
DB 360 TTTAAATCAAGATGAGTATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCT 419
QY 101 MetThrValIleSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
   |||
DB 420 ATACCGGTGTCAGTTTCATTCAGTGGATTAATCCTTTGACCGGAATGTGTATCCAG 479
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
   |||
DB 480 CTGTTAAATGAATGACGAGAGATGCTGCACCAATCATTTAGCGCCACTCATGCGCAAG 539
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
   |||
DB 540 TCACATGACGCGGTTAATATATGTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 599
QY 161 LeuTyraAspProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
   |||
DB 600 TTGTATTAATCTTTTGGGAATTTTAATACCCACTTACAAAACCTATGTGATGATGATAC 659
QY 181 LysMetLeuAspGluGluAsnIle 188
   |||
DB 660 AAAATGTTGATGACAGAACATATA 683
DB
RESULT 11
ADA11570
ID ADA11570 standard; DNA; 1943 BP.
XX
XX ADA11570;
AC
XX
XX 06-NOV-2003 (first entry)
DT
XX
XX Human cDNA encoding a novel secreted protein, SEQ ID NO 98.
DE
XX
XX Cancer; inflammation; immune disorder; neurological disorder;
KM blood clotting disorder; food additive; food preservative;
KW storage capability; fat content; nutritional component; ds; gene; human.
XX
OS Homo sapiens.
XX
XX US2003055236-A1.
PN
XX
XX 20-MAR-2003.
PD
XX
XX 14-MAR-2002; 2002US-00097065.
PF
XX
XX 18-DEC-1997; 97US-0068006P.
XX
XX 18-DEC-1997; 97US-0068007P.
PR 18-DEC-1997; 97US-0068008P.
PR 18-DEC-1997; 97US-0068003P.
PR 18-DEC-1997; 97US-0068054P.
PR 18-DEC-1997; 97US-0068057P.
PR 18-DEC-1997; 97US-0068064P.
PR 18-DEC-1997; 97US-0070923P.
PR 19-DEC-1997; 97US-0068169P.
PR 19-DEC-1997; 97US-0068365P.
PR 19-DEC-1997; 97US-0068367P.
PR 19-DEC-1997; 97US-0068368P.
PR 19-DEC-1997; 97US-0068369P.
PR 17-DEC-1998; 98MO-US027059.
PR 17-JUN-1999; 99US-00334595.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Moore PA, Ruben SM, Carter KC, Shi Y, Rosen CA, Soppet DR;
PI Kyaw H, Wei Y, Florence KA, Duan DR, Florence C, Greene JM, Feng P;
PI Ferré AM, Yu G, Janat F, Ni J;
XX
XX WPI; 2003-567105/53.

```

```

DR P-PSDB; ADA11694.
XX
XX New secreted HKAR24 nucleic acid molecules and polypeptides, useful for
PT preventing treating, or ameliorating a medical condition, such as
PT cancer, inflammation, immune disorders, neurological and blood clotting
PT disorders.
XX
XX PS. Claim 1; SEQ ID NO 98; 118pp; English.
XX
XX The invention relates to an isolated HKAR24 nucleic acid molecule. The
CC polypeptides, nucleic acids and antibodies are useful for diagnosing a
CC pathological condition or a susceptibility to a pathological condition,
CC for preventing, treating, or ameliorating a medical condition, such as
CC cancer, inflammation and other immune disorders, neurological and blood
CC clotting disorders. The nucleic acids are also useful for chromosome
CC identification, radiation hybrid mapping or long-range restriction
CC mapping. The polypeptides and antibodies are useful for providing
CC immunohistochemistry assays. The polypeptide, polynucleotide, agonist or
CC antagonist may also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content or other
CC nutritional components. The present sequence represents cDNA encoding a
CC novel human secreted protein. Note: The sequence data for this patent did
CC not form part of the printed specification but was obtained in electronic
CC format directly from USPTO at
CC seqdata.uspto.gov.uk/sequence.html?DocID=20030055236.
XX
XX SO Sequence 1943 BP; 629 A; 325 C; 399 G; 586 T; 0 U; 4 Other;

Alignment Scores:
Pred. No.: 2,06e-97 Length: 1943
Score: 948.00 Matches: 186
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 2
Query Match: 98.54% Indels: 0
DB: Gaps: 0

US-10-627-571-2 (1-188) x ADA11570 (1-1943)
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAAGlnLysIleLeu 20
   |||
DB 120 ATGGCCACAGATGTCCTTAATTCCAAAAACCTGGCGCTTCANGCAAAAAGAAATCTTG 179
QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
   |||
DB 180 GGTAAATGTGTCTCAATTCATCCGACACCTTATATGACGACAAAGTGTAGGTG 239
QY 41 LeuAspGluLeuTyraArgValThrArgGluTyrrThrGlnAsnLysGluAAGluLys 60
   |||
DB 240 CTGATGAGACTCTACAGAGTGACAGGGAGTACACCAAAACAAAGAGGACAGGAAG 299
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnGln 80
   |||
DB 300 ATCATCAAGAACCTTCAAGACAGTCATCAAGCTTCCATCTTTATAGATATATCCG 359
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAla 100
   |||
DB 360 TTTAAATCAAGATGAGTATGATGAGAAATTTAAGAAAGATTTCATCAGCTTGCT 419
QY 101 MetThrValIleSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
   |||
DB 420 ATACCGGTGTCAGTTTCATTCAGTGGATTAATCCTTTGACCGGAATGTGTATCCAG 479
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
   |||
DB 480 CTGTTAAATGAATGACGAGAGATGCTGCACCAATCATTTAGCGCCACTCATGCGCAAG 539
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
   |||
DB 540 TCACATGACGCGGTTAATATATGTGTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 599
QY 161 LeuTyraAspProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
   |||

```

Db 600 TTGATAATCCCTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATAC 659
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 660 AAAATGTTGATGATGAAGAACAATA 683

RESULT 12
 ADA56241
 ID ADA56241 standard; DNA; 1943 BP.
 AC ADA56241;
 XX
 XX
 DT 20-NOV-2003 (first entry)
 DE Gene encoding human secreted protein #420.
 XX
 XX
 XX Immunosuppressive; antiinflammatory; antiaesthetic; antiallergic;
 KM cytosolic; cerebroprotective; neuroprotective; nootropic;
 KM cardiovascular; antiarteriosclerotic; gene therapy;
 KM human secreted protein; immune disorder; inflammation;
 KM respiratory disorder; cancer; CNS disorder; neurodegenerative disorders;
 KM inflammatory bowel disease; nephritis; Crohn's disease; asthma; allergy;
 KM multiple sclerosis; ischaemic brain injury; Parkinson's disease;
 KM Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
 KM triple helix formation; antisense gene therapy; forensic biology; ds;
 KM gene.
 XX
 OS Homo sapiens.
 XX
 PN WC002102994-A2.
 XX
 XX 27-DEC-2002.
 XX
 PF 19-MAR-2002; 2002WC-US008278.
 XX
 PR 21-MAR-2001; 2001US-0277340P.
 PR 19-JUL-2001; 2001US-0306171P.
 PR 13-NOV-2001; 2001US-0331287P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Ruben SM;
 XX
 DR WPI; 2003-167512/16.
 DR P-PSDB; ADA57137.
 PT
 PT New human secreted polypeptides and polynucleotides, useful for
 PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
 PT conditions, respiratory disorders, cancers, CNS disorders, or
 PT neurodegenerative disorders.
 XX
 XX
 PS Claim 21; SEQ ID NO 430; 1754bp; English.
 XX
 XX The invention relates to 592 new human secreted polypeptides useful for
 CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
 CC conditions, respiratory disorders, cancers, CNS disorders, or
 CC neurodegenerative disorders, or polypeptides comprising an amino acid
 CC sequence at least 95% identical to the new sequences. The polypeptides,
 CC antibodies or antibody fragments that bind to the polypeptides, nucleic
 CC acids encoding the polypeptides, agonists or antagonists that binds to
 CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
 CC compositions for diagnosing, treating or preventing an e.g. immune
 CC disorders, inflammatory conditions (e.g. inflammatory bowel disease,
 CC nephritis or Crohn's disease), respiratory disorders (e.g. asthma and
 CC allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders
 CC (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative
 CC disorders (e.g. Parkinson's disease or Alzheimer's disease), and
 CC cardiovascular disorders (e.g. atherosclerosis or myocarditis). The
 CC polynucleotides are useful for chromosome identification, chromosome
 CC mapping, for controlling gene expression through triple helix formation
 CC or antisense DNA or RNA, in gene therapy, for identifying individuals
 CC from minute biological samples, in forensic biology, and as hybridization
 CC probes. The polypeptides are useful for as molecular weight markers on

CC sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)
 CC gels, to raise antibodies, for testing biological activities, and for
 CC treating or preventing neural disorders, immune system disorders,
 CC muscular, reproductive, gastrointestinal, pulmonary, cardiovascular,
 CC renal, proliferative and/or cancerous diseases. This sequence corresponds
 CC to a gene encoding one of the polypeptide of the invention. Note: The
 CC sequence data for this patent did form part of the printed specification,
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX
 SQ Sequence 1943 BP; 629 A; 325 C; 399 G; 586 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 2,06e-97 Length: 1943
 Score: 948.00 Matches: 186
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 2
 Query Match: 98.54% Indels: 0
 DB: 10 Gaps: 0

US-10-627-571-2 (1-188) x ADA56241 (1-1943)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAgiLysLysIleLeu 20
 Db 120 ATGGCCACAGATGCTTTTAATTCAAAACCTGCGCTTCANGCACAAAAGAAATCTTG 179
 QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 180 GGTAATAATGGTGTGCAAAATCCATCGCCACACCTTATAGACACAAAGTAAGTGGTGTG 239
 QY 41 LeuAspGluLeuThrArgValThrArgGluThrGlnAsnLysLysGluAgiLys 60
 Db 240 CTGGATGAGCTCTACAGAGTGAACAGAGGTACACCCAAACAAAGAGAGCGAGAAAG 299
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuThrArgAsnAsnGln 80
 Db 300 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGCGCAATCTTTATAGAAATATCG 359
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 360 TTTAATCAAGATGAGCTACATTGATGAGAAATTTAAGAAGAAATTCATCATCGCTTGGCT 419
 QY 101 MetThrValIleSerPheHisGlnValAspThrThrPheAspArgAsnValLeuSerArg 120
 Db 420 ATGACCGGTGCTGATTCATCATGAGTGAATTAACCTTGACCGGATGTGTTATCCAGG 479
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 480 CTGTAAATGAATGACAGAGAGATGCTGCACAAATATTTCAGGCCACTGCTGCACAG 539
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
 Db 540 TCACATGACGACGGTTAATAATGTTTGAATCTTTTCGATTTGTAATTTTGGCTGCC 599
 QY 161 LeuThrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 600 TTGTATAATCCCTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATCAAC 659
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 660 AAAATGTTGATGATGAAGAACAATA 683

RESULT 13
 ACN44177
 ID ACN44177 standard; cDNA; 2087 BP.
 XX
 XX ACN44177;
 AC
 DT 18-NOV-2004 (first entry)
 XX
 XX Mouse mRNA sequence mCT17722.
 DE
 XX
 KM Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ds.

```

XX OS Mus musculus.
XX PN MO2003073826-A2.
XX PD 12-SEP-2003.
XX PF 28-FEB-2003; 2003WO-US006235.
XX PR 01-MAR-2002; 2002US-00087192.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX PT WPI; 2003-328604/31.
XX PS Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX PT comprises a nucleotide sequence.
XX PS Claim 1; SEQ ID NO 494; Opp; English.
XX CC The present invention relates to novel DNA and protein sequences which
XX CC are associated with carcinomas. The sequences are useful for: (i) for
XX CC screening drug candidates; (ii) for screening of bioactive agent capable
XX CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX CC a bioactive agent capable of modulating the activity of CAP; (iv) for
XX CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX CC determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX CC carcinoma including lymphoma. The present sequence is one such CA gene
XX CC sequence. Note: This patent is an equivalent to basic patent
XX CC US2002182586A1, for which no sequence data was published
XX CC
SQ Sequence 2087 BP; 556 A; 487 C; 508 G; 536 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,4e-94 Length: 2087
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 11 Gaps: 0

US-10-627-571-2 (1-188) x ACN44177 (1-2087)
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeu 20
DB 507 GGGCTACAGATGCTTCAATTCACAAAACCTGGCCGTTACAGCAAAAGAGATCTCG 566
QY 21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 567 GGCAAATATGTATCCAAATCCATCGCCACCAACGCTGATGACACACGACGAGCGGTG 626
QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGluLys 60
DB 627 CTAGATGAGCTGTACAGGGGTGACCAAGAGTACACCCAGAAACAAAGAGGCGGAGAGAG 686
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
DB 687 GTCATCAAGAACTCTATCAAGACGGTCAATCAAGCTGGCGTCTCCACAGAACTCAAG 746
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysLysValIleGlnLeuAla 100
DB 747 TTCATCAAGACGAGCTGGCGCTCATGAGAAAGTTCAAGAAAGAGTGCACCGCTTGGC 806
QY 101 MetThrValValSerPheHISGlnValaLeuPyrTyrThrPheAspArgAsnValLeuSerArg 120
DB 807 ATGACGGTGTGACGCTTCCACCAAGTAAAGTACACCTTCGACCGCAATGTGCTCAGG 866

```

```

QY 121 LeuLeuAsnGluCysArgGluMetLeuHISGlnIleIleGlnArgHISLeuThrAlaLys 140
DB 867 CTGCTGAACGATGTCGAGAGCTCTTACACGAGATCATTCACGCGCACCTTACCGCCAG 926
QY 141 SerHISGlyArgValAsnAsnValPheAspHISpheSerAspCysGluPheLeuAlaAla 160
DB 927 TCTCAGCAGCGGGTTAATATGTCCTTTGACCATTTTTCAGATTGATTTTGGCTGCC 986
QY 161 LeuTyrAsnProPheGlyAsnPhelyProHISLeuGlnLysLeuCysAspGlyIleAsn 180
DB 987 TTGTACAAATCCCTTGGAAAGTTTAAACCTCACTTACAGAACTTGGCAGCGCATCAAC 1046
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 1047 AAAATGTTGATGAAGAGAACATA 1070

RESULT 14
ACN44176
ID ACN44176 strand; DNA; 62231 BP.
XX ACN44176;
XX ACN44176;
XX DT 18-NOV-2004 (first entry)
XX DE Mouse genomic sequence MCG15520.
XX KW Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX OS Mus musculus.
XX PN MO2003073826-A2.
XX PD 12-SEP-2003.
XX PF 28-FEB-2003; 2003WO-US006235.
XX PR 01-MAR-2002; 2002US-00087192.
XX PA (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX PT WPI; 2003-328604/31.
XX PS Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX PT comprises a nucleotide sequence.
XX PS Claim 1; SEQ ID NO 493; Opp; English.
XX CC The present invention relates to novel DNA and protein sequences which
XX CC are associated with carcinomas. The sequences are useful for: (i) for
XX CC screening drug candidates; (ii) for screening of bioactive agent capable
XX CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX CC a bioactive agent capable of modulating the activity of CAP; (iv) for
XX CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX CC determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX CC carcinoma including lymphoma. The present sequence is one such CA coding
XX CC sequence. Note: This patent is an equivalent to basic patent
XX CC US2002182586A1, for which no sequence data was published
XX CC
SQ Sequence 62231 BP; 15482 A; 12591 C; 12729 G; 16514 T; 0 U; 4915 Other;

Alignment Scores:
Pred. No.: 2,97e-92 Length: 62231
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 11 Gaps: 0

```


US-10-627-571-2 (1-188) x ACN44176 (1-62231)

```

QY      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
      : : : : :
DB      50651 GTGGCTCAGAGTGTCTTCATTCATAAATCCAAAACCTGGCCGTTGAGGACAAAGAAAGATCCTG 50710

QY      21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      : : : : :
DB      50711 GGCAAAATGGATCAATCAATCCATCGCCACCGCTGATCGACGACACAGAGAGAGGTG 50770

QY      41 LeuAspGluLeuTyraArgValThrArgGluTyrrThrGlnAsnLysLysGlnAlaGluLys 60
      : : : : :
DB      50771 CTGATGATAGCTGTATCAGGGTGCACCAAGAGATACCCAGAAACAAAGAGAGGCGGAGAGG 50830

QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnAsnGln 80
      : : : : :
DB      50831 GTCATCAAGAACCCATCAAGACGCGTATCAAGCTGGCCGCTCCACAGGAACAATCAG 50890

QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla 100
      : : : : :
DB      50891 TTCATCAAGACGAGCTGGCGCTCATGAGAAAGTTCAAGAAAGAGTGCCACGACTTGCC 50950

QY      101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValIleuSerArg 120
      : : : : :
DB      50951 ATGACGCTCGTCACTCCACCAAGTATGAGTACACCTTCGACCCCAATGGTGTCCAGG 51010

QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      : : : : :
DB      51011 CTGCTGAACGAGTCCGAGAGCTCTACACAGATCATTTACAGGCCCATTTACCGCCAG 51070

QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
      : : : : :
DB      51071 TCTCAGCAGCGGGTAAATTAATGCTTTGACCAATTTGATGTGATTTTGGCTGCC 51130

QY      161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
      : : : : :
DB      51131 TTGTACATACCTCTTTGAAAAGTTTAAACCTCACTTACGAAACCTTTCGAGGCGATAC 51190

QY      181 LysMetLeuAspGluGluAsnIle 188
      : : : : :
DB      51191 AAAATGTTGATGAAGAAACATTA 51214

RESULT 15
AAH07403
ID      AAH07403 standard; cDNA; 816 BP.
XX
AC      AAH07403;
XX
DT      26-JUN-2001 (first entry)
XX
DE      Human cDNA clone (5'-primer) SEQ ID NO:4238.
XX
KM      Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS      Homo sapiens.
XX
PN      BP1074617-A2.
XX
PD      07-FEB-2001.
XX
PF      28-JUL-2000; 2000EP-00116126.
XX
PR      29-JUL-1999; 99JP-00248036.
XX
PR      27-AUG-1999; 99JP-00300253.
XX
PR      11-JAN-2000; 2000JP-00118776.
XX
PR      02-MAY-2000; 2000JP-00183767.
XX
PR      09-JUN-2000; 2000JP-00241899.
XX
(HELI-) HELIX RES INST.
PA
XX      Oka T, Isogai T, Nishikawa T, Hayashi K, Satou K, Yamamoto J,
PI      Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

```

```

XX      DR      WPI; 2001-318749/34.
XX
XX      PT      Primer sets for synthesizing polynucleotides, particularly the 5602 full-
PT      length cDNAs defined in the specification, and for the detection and/or
PT      diagnosis of the abnormality of the proteins encoded by the full-length
PT      cDNAs.
XX
XX      Claim 1; SEQ ID NO 4238; 2537bp + Sequence Listing; English.
XX
XX      The present invention describes primer sets for synthesizing 5602 full-
CC      length cDNAs defined in the specification. Where a primer set comprises:
CC      (a) an oligo-dT primer and an oligonucleotide complementary to the
CC      complementary strand of a polynucleotide which comprises one of the 5602
CC      nucleotide sequences defined in the specification; where the
CC      oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC      of an oligonucleotide comprising a sequence complementary to the
CC      complementary strand of a polynucleotide which comprises a 5'-end
CC      sequence and an oligonucleotide comprising a sequence complementary to a
CC      polynucleotide which comprises a 3'-end sequence; where the
CC      oligonucleotide comprises at least 15 nucleotides and the combination of
CC      the 5'-end sequence/3'-end sequence is selected from those defined in the
CC      specification. The primer sets can be used in antisense therapy and in
CC      gene therapy. The primers are useful for synthesizing polynucleotides,
CC      particularly full-length cDNAs. The primers are also useful for the
CC      detection and/or diagnosis of the abnormality of the proteins encoded by
CC      the full-length cDNAs. The primers allow obtaining of the full-length
CC      cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC      AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
CC      represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC      oligonucleotides, all of which are used in the exemplification of the
CC      present invention
XX
XX      SQ      Sequence 816 BP; 262 A; 155 C; 180 G; 212 T; 0 U; 7 Other;
XX
XX      Alignment Scores:
XX      Pred. No.: 2,816-94 Length: 816
XX      Score: 916.00 Matches: 184
XX      Percent Similarity: 97.89% Conservative: 2
XX      Best Local Similarity: 96.84% Mismatches: 2
XX      Query Match: 95.22% Indels: 2
XX      DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x AAH07403 (1-816)
QY      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
      : : : : :
DB      101 GTGGCCACAGATGTCTTAAATTCCAAAAACCTGGCCGTTGAGGACAAAGAAAGATCTTG 160

QY      21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      : : : : :
DB      161 GGTAATAATGGTGTCCAAATCCATCGCCACCACTTATATAGACACAAGTATGAGGTG 220

QY      41 LeuAspGluLeuTyraArgValThrArgGluTyrrThrGlnAsnLysLysGlnAlaGluLys 60
      : : : : :
DB      221 CTGGATGACCTTCACAGAGTGCACAGGAGTACACCAACCAAGAGAGGAGGAGAG 280

QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnAsnGln 80
      : : : : :
DB      281 ATCATCAAGAACCTTCATCAAGACGATCATCAAGCTGCCATTTTATAGAAATATACG 340

QY      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla 100
      : : : : :
DB      341 TTTAATCAAGATGAGTATGATGAGGAATTTAAGAAAGAAATTCATCAGCTTGCT 400

QY      101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValIleuSerArg 120
      : : : : :
DB      401 ATGACCGTGTGATTCATCAAGGATGATTAACCTTTGACCGGAAATGTATTCAGG 460

QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      : : : : :
DB      461 CTGTTAAATGAATGACGAGAGATGCTGCACCAATATCTCAGCGCACCTCACTGCGCAAG 520

```


QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla160
|||
DB 521 TCACATGACGCGGTAAATATGCTTGCATCATTTTCAGATTGTGAATTTGGCTGCC 580
QY 161 LeuTyrAsnProPhGlyAsnPhelysProHisLeuGlnIlyLeuCysAspGlyIleAsn 180
581 TTGTATTAATCCCTTTGGGAATTTTAAACCCCACTTCAAAAATAAGTGTATCATCAAC 640
QY 181 LysMet--LeuAspGluGluAsnIle 188
|||
DB 641 AAAATTGTTGGATGGAAGAACATATA 666
RESULT 16
ACH89719
ID ACH89719 standard; DNA; 544 BP.
XX ACH89719;
XX
AC ACH89719;
XX
DT 29-JUL-2004 (first entry)
XX
XX Human genome derived single exon probe #22914.
DE Human genome derived single exon probe #22914.
XX
KW Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
XX US2003194704-A1.
PN 16-OCT-2003.
XX
XX 03-APR-2002; 2002US-00029386.
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
PS Claim 1; SEQ ID NO 22914; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
XX expression, comprising any of the 27,400 fully defined nucleotide
XX sequences in the specification, or their complements or fragments, and
XX encoding at least 8 amino acids of any of the 6888 amino acid sequences
XX fully defined in the specification. The probe is a single exon probe that
XX hybridizes under high stringency conditions to a nucleic acid molecule
XX expressed in human cells or tissues. Also included are a spatially-
XX addressable set of single exon nucleic acid probes for measuring human
XX gene expression (comprising a plurality of single exon nucleic acid
XX probes cited above, where each of the plurality of probes is separately
XX and addressably isolatable or amplifiable from the plurality), a single
XX exon microarray for measuring human gene expression, a method of
XX measuring human gene expression, a vector comprising the single exon
XX probe cited above, an ORF-encoded peptide comprising at least 8
XX contiguous amino acids of any of the above-mentioned amino acid
XX sequences (optionally with conservative amino acid substitutions), an
XX isolated antibody that binds specifically to a peptide cited above,
XX methods of selling and/or licensing single exon probes or microarrays to
XX a customer desiring to measure gene expression, a method of providing
XX human gene expression data by a subscription, and a computer-readable
XX storage medium which contains a database having a plurality of records
XX (each record including data on the expression of a single exon probe
XX cited above. The probe, methods and apparatus are useful in gene

CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterizing
CC alternative splicing events, in detecting and characterizing gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docID=20030194704
XX
SQ Sequence 544 BP; 175 A; 109 C; 117 G; 143 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2,78e-94 Length: 544
Score: 914.00 Matches: 179
Percent Similarity: 99.44% Conservative: 0
Best Local Similarity: 99.44% Mismatches: 1
Query Match: 95.01% Indels: 0
DB: 12 Gaps: 0
US-10-627-571-2 (1-188) x ACH89719 (1-544)
QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeuGly 21
|||
DB 3 GCCACAGATGCTTTTAAATCCAAAACCTGGCCGCTTCAGGCCCAAAAAGAGATCTTGGGT 62
QY 22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
63 AAAATGCTGTCCAAATTCATCGCCACACCTTAATATAGACACAAAGTAGTAGGCTGCTG 122
QY 42 AspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGluLysLys 61
123 GATGAGCTCTTACAGAGTGCACAGGAGTACACCCAAAACMAAGAGAGCAAGAGATGC 182
QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGlnPhe 81
183 ATCAAGAACCTCATCAAGACAGTCAATCAAGCTGGCCATTTCTTTATAGAAATATCACTTT 242
QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
243 AATCAAGATGAGCTAGCATTTGATGGAAATTTAAGAAAGATTATCATCAGCTTGCTATG 302
QY 102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
303 ACCGTGCTCAGTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTGTATCCAGGCTG 362
QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
363 TTAATATGAATGACAGAGAGTGTGCACCAATATTCAGCGGCACCTCATCCAGATGCA 422
QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
423 CATGACGCGGTAAATATATGTTGATCATTTTTCAGATGTGGAATTTTGGCTGCTTG 482
QY 162 TyrAsnProPhGlyAsnPhelysProHisLeuGlnIlyLeuCysAspGlyIleAsnLys 181
483 TATATATCCCTTTGGGAATTTTAAACCCCACTTCAAAAATAAGTGTATCATCAACAA 542
RESULT 17
AD063108
ID AD063108 standard; cDNA; 1986 BP.
XX
XX AD063108;
XX
AC AD063108;
XX
DT 07-OCT-2004 (first entry)
XX
XX Novel human cDNA sequence #269.
XX
XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;

KM neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
 KM cancer.
 OS Homo sapiens.
 PN EP1440981-A2.
 XX
 PD 28-JUL-2004.
 XX
 PF 21-JAN-2004; 2004EP-00001196.
 XX
 PR 21-JAN-2003; 2003JP-00102206.
 PR 09-MAY-2003; 2003JP-00131392.
 XX
 PA (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX
 PI Isegai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Nagai K, Irie R;
 XX WPI; 2004-535376/52.
 DR P-PSDB; ADQ65296.
 XX
 PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX
 PS Claim 1; SEQ ID NO 269; 2449pp; English.
 XX
 CC The invention relates to 2495 novel polynucleotides (I) and their encoded
 CC polypeptides, sequences hybridizing to these nucleotides, sequences
 CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a nucleotide
 CC sequence of the invention.
 XX
 SQ Sequence 1986 BP; 536 A; 458 C; 497 G; 495 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 1.08e-63 Length: 1986
 Score: 651.00 Matches: 114
 Percent Similarity: 89.07% Conservative: 49
 Best Local Similarity: 62.30% Mismatches: 20
 Query Match: 67.67% Indels: 0
 DB: Gaps: 0
 US-10-627-571-2 (1-188) x ADQ63108 (1-1986)
 QY 4 AspValPheAsnSerIleAsnLeuAlaValGlnAlaGlnIleuGlyMet 23
 DB 152 GATGTTTTCAGTTCAGGAGTTCCTTCCCTTCAAGCCAGAAAGATTCTGACCAATA 211
 QY 24 ValSerIleSerIleAlaThrThrIleuIleAspThrSerGluValIleAspGlu 43
 DB 212 GCCAGCAAACTGGGCAACATGTTGATGATGACCCAGCGGATCTTGAGAG 271
 QY 44 LeuTyrArgValThrArgGluTyrThrGlnAsnIleuGlnIleuGlyMet 63
 DB 272 CTCTACAAATCTCACAAGACACACACACACACACACACACACACACACAC 331
 QY 64 AsnIleuIleuThrValIleuLeuAlaIleuTyrArgAsnGlnIleuAsn 83
 DB 332 GACTTATCAAGGTGGCGATCAAAATCGGATCCTTACCGGAACAACGATTAGCAA 391
 QY 84 AspGluLeuAlaLeuMetGluValPheIleuGlyValIleuGlnIleuAlaMetThrVal 103
 DB 392 GAGAGCTGTTATGAGAGAGTTCCGAGAGAGCTGAACACGACCGCATGACCAT 451
 QY 104 ValSerPheIleuValAspTyrThrPheAspArgAsnValIleuSerArgLeuAsn 123
 DB 452 GTACAGTTTATAGAGGTGAATACACCTTGATGAGAACGTGCTCTCATCTCTGCAT 511

QY 124 GluCyArgGluMetLeuHsiGlnIleIleGlnAlaArgHsiLeuThrAlaIleSerHsiGly 143
 DB 512 GAGTGAAGGACCTGGGCTGAGTGAACCTGACCGGACCTGACCGGACCGG 571
 QY 144 ArgValAsnAsnValPheAspHsiPheSerAspCysGluPheLeuAlaIleuTyrAsn 163
 DB 572 CGCATCAACACCTCTTAAACACCTTCCGAGTGTGAGTCTCTCTCCACCTCTAAGT 631
 QY 164 ProPheGlyAsnDheliysProHsiLeuGlnIleuCyAspGlyIleAsnIleuMetLeu 183
 DB 632 CTGATGAGAGACTGTAGCCCACTCAAGAGATTTGTGAAGAAATCAATTAAGTTCTA 691
 QY 184 AspGluGlu 186
 DB 692 GATGAGAA 700
 RESULT 18
 AD30769
 ID AD30769 standard; cDNA; 969 BP.
 AC
 XX AD30769;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel cDNA sequence, SEQ ID NO:851.
 XX
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antihaemic; anticoagulant; thrombolytic; vulnerary;
 KW anticancer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 15q21.2; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO2003029271-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002WO-US030474.
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac R;
 DR WPI; 2003-371981/35.
 DR P-PSDB; AD31740.
 XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Claim 1; SEQ ID NO 851; 1185pp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (AD329919-
 CC AD30889) and the polypeptides they encode (AD30890-AD31860). The
 CC invention also relates to nucleic acid sequences over 9% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the

CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridization probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 969 BP; 264 A; 261 C; 257 G; 187 T; 0 U; 0 Other;

XX Alignment Scores:
 XX Pred. No.: 7,45e-63 Length: 969
 XX Score: 640.00 Matches: 112
 XX Percent Similarity: 88.46% Conservative: 49
 XX Best Local Similarity: 61.54% Mismatches: 21
 XX Query Match: 66.53% Indels: 0
 XX DB: Gaps: 0

US-10-627-571-2 (1-188) x ADC30769 (1-969)

QY 4 AsgValPheasnserysaenleuAlaValGlnAlaGlnlysllyleuGlylyMet 23
 DB |||||.....
 DB 322 GAGTGTTCAGTTCAGGATCTTCGCTTCAGCCAGAGAAATTCGACCAAAATA 381
 QY 24 ValserlyserlleThrlThrleuIleAspAspThrsersergluValLeuAsglu 43
 DB |||||.....
 DB 382 GCCAGCAAACTGTGGCCAAACATGTTGATGACACCGACGAGATCTTGTATGAG 441
 QY 44 LeuTyArgValThrArgIuTyTrhArgInAsnlyslsGluIaGluyslylleys 63
 DB |||||.....
 DB 442 CTCACAAAGTCACCAAGGACACACACACAAAGAGAAAGCCCAAGATCAGTAA 501
 QY 64 AsnleuIlelyThrValIlelyLeuAlaIleleuTyArgAsnAsngInPheAsnGln 83
 DB |||||.....
 DB 502 GACTTATCAAGGTGGCGATCAAAATCGGATCCTCTACCGGAAACACAGTTAGCCAA 561
 QY 84 AsgGluLeuAlaLeuMetGluysPheIyValyValIleGlnleuAlaMetTrhVal 103
 DB |||||.....
 DB 562 GAGGAGCTGTATTGTGGAGAAAGTTCGGAAAGAGCTGAACACAGCCGATGACCAT 621
 QY 104 ValserPhehisGlnValAspTyTrhPheAspArgAsnValleuserArgLeuLeuAsn 123
 DB |||||.....
 DB 622 GTGAGCTTCATGAGGTGGAATACACTTGCATAGGAAGTGTCTCCATCTCCGCA 681
 QY 124 GluCyArgGluMetleuHisGlnIleleGlnArgHisleuThralaIysSerHisGly 143
 DB |||||.....
 DB 682 GAGTCGAAGGACCTGGTGATGAACTGTGACCGGACCTGACGCGCCAGAGACCCAGCG 741
 QY 144 ArgValAsnAsnValPheAspHisPheSerAspCyGluPheleuAlaIaIeUuTyArg 163
 DB |||||.....
 DB 742 CGCATCAACACGCTTTAAACACTTGGAGATGGAGTTCCTCTCCACCTCTATAGT 801
 QY 164 ProPheGlyAsnAsnleuysProHisleuGlnlyleuCyAspGlyIleAsnlyMetLeu 183
 DB |||||.....
 DB 802 CTGATGAGGAGACTGTAGGCGCAACCTCAAGAGATTTGTAAAGAAATCATTAAGTTGCT 861
 QY 184 AspGlu 185
 DB |||||.....
 DB 862 TCCAG 867

RESULT 19
 ACH92216
 ID ACH92216 standard; DNA; 527 BP.
 XX
 AC ACH92216;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Human genome derived single exon probe #25411.
 XX
 KW Human; probe; ss; gene expression; single exon probe; microarray;
 KW alternative splicing event; genomic alteration.
 XX
 OS Homo sapiens.
 XX
 PN US2003194704-A1.
 XX
 PD 16-OCT-2003.
 XX
 PF 03-APR-2002; 2002US-00029386.
 XX
 PR 03-APR-2002; 2002US-00029386.
 XX
 PA (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 PI Penn SG, Rank DR, Hanzel DK;
 XX
 DR WPI; 2004-119264/12.
 XX
 PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 PS
 PS Claim 1; SEQ ID NO 25411; 80bp; English.
 XX
 CC The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressed set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterizing
 CC alternative splicing events, in detecting and characterizing gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20030194704

```

XX  SQ      Sequence 527 BP; 163 A; 129 C; 128 G; 107 T; 0 U; 0 Other;
XX
XX  Alignment Scores:
XX      Pred. No.:      8,49e-60      Length:      527
XX      Score:          610.00      Matches:      106
XX      Percent Similarity: 88.37%      Conservative: 46
XX      Best Local Similarity: 61.63%      Mismatches: 20
XX      Query Match:      63.41%      Indels:      0
XX      DB:              12      Gaps:      0
XX
XX  US-10-627-571-2 (1-188) x ACH92216 (1-527)
XX
QY  15  AAGlnlylylylyleleuglylymecvalserlyserllealatrthrlleuileasp 34
DB  3  GCCCAGAGAAAGATTCTGCAGCAAAATAGCCAGAAACTGTGGCCAACTGATTGATTGAT 62
QY  35  AspThrsersergluvalleuaspgluleuTyArgValThrArgglutyrrthrglnasn 54
DB  63  GACACCGACGACGAGATCTTTGATGAGCTCTCAAAAGTCAACAAAGGACACACACAC 122
QY  55  LyelysgluAaIaglylylylelyasnlleuilelythrVallelyleuAaile 74
DB  123  AAGAGAGAACCCACAAAGATCATGAAGACTTATATCAAGGTGGGATCAAAATCGGGATC 182
QY  75  LeuTyArgAsnAsnnglnPheAsnnglnAspGluLeuAaileuWecGulysPheLylyls 94
DB  183  CTCTACCGGAAACAACCGATTAGCCAAAGAGCTGTGTATTGTGGAGAAATTCCGGAG 242
QY  95  LyeValHlaglIneulAmectThrValSerPheHlglInValAspTyrrthrrpheasp 114
DB  243  AAGGTGAACCAAGACCGCATGACCATTTGCTATATAGAGTGGAAATACACCTTGAT 302
QY  115  ArgAsnValleuSerArgleuLeuAsnGluCysArggluWecLeuHlglInleleGln 134
DB  303  AGGAGCGTGTCTTCATCTCTCGCAGATGACGAGAGACTGTGGCATGAACGTGTCCAG 362
QY  135  ArgHlglLeuThralAlySerHlglYArgValAsnAsnValPheAspHlspheSerAsp 154
DB  363  CGGCACCTGACGCCGACCCAGCCGCGCATCAACACACTTTTAACTTGGCCGAT 422
QY  155  CysGluPheleuAaIaIeutyrrAsnProPheGlyAsnheLySProhlsleuGlnlys 174
DB  423  GTGGAGTTCTCTCCACCTCTATAGTCTGATGAGACGTGAGCCCAACCTCAAGAG 482
QY  175  LeuCyAspGlylleAsnlysmecleuAspGlu 186
DB  483  ATTGTCAAGGATCAATTAAGTGTCTGATGAGAAA 518
XX
XX  RESULT 20
XX  ABS78742 standard; DNA: 619 BP.
XX
XX  ABS78742:
XX
XX  16-DEC-2002 (first entry)
XX
XX  DNA encoding human NOVX17b protein.
XX
XX  Human; NOVX; human disease; NOVX-associated disorder; cancer; addiction;
XX  Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
XX  tuberculous sclerosis; hypercalcaemia; Parkinson's disease; depression;
XX  Huntington's disease; cerebral palsy; epilepsy; Lesch-Nyhan syndrome;
XX  multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;
XX  obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
XX  infertility; inflammatory bowel disease; atherosclerosis; hypertension;
XX  scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
XX  acthma; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
XX  bacterial infection; parasitic infection; graft-versus-host disease;
XX  cell differentiation; cell proliferation; haematopoiesis; wound healing;
XX  angiogenesis; gene; ds.
XX
XX  Homo sapiens.

```

```

XX  PN      WO200272770-A2.
XX  PD      19-SEP-2002.
XX  PF      08-MAR-2002; 2002WO-US007283.
XX  PR      08-MAR-2001; 2001US-0274281P.
XX  PR      09-MAR-2001; 2001US-0274849P.
XX  PR      12-MAR-2001; 2001US-0275235P.
XX  PR      13-MAR-2001; 2001US-0275579P.
XX  PR      13-MAR-2001; 2001US-0275601P.
XX  PR      14-MAR-2001; 2001US-0276000P.
XX  PR      20-MAR-2001; 2001US-0277239P.
XX  PR      20-MAR-2001; 2001US-0277327P.
XX  PR      20-MAR-2001; 2001US-0277328P.
XX  PR      21-MAR-2001; 2001US-0277791P.
XX  PR      22-MAR-2001; 2001US-0277833P.
XX  PR      23-MAR-2001; 2001US-0278152P.
XX  PR      26-MAR-2001; 2001US-0278894P.
XX  PR      27-MAR-2001; 2001US-0279036P.
XX  PR      28-MAR-2001; 2001US-0279344P.
XX  PR      30-MAR-2001; 2001US-0280233P.
XX  PR      02-APR-2001; 2001US-0280802P.
XX  PR      02-MAY-2001; 2001US-0288148P.
XX  PR      31-MAY-2001; 2001US-0294821P.
XX  PR      31-OCT-2001; 2001US-0335302P.
XX  PR      04-DEC-2001; 2001US-0338375P.
XX  PR      07-MAR-2002; 2002US-00094466.
XX
XX  (CURA-) CURAGEN CORP.
XX
XX  Spytek KA, Vernet CA, Tchertnev VT, Malyankar UM, Gerlach VL;
XX  Li L, Zernusen BD, Paturajan M, Gusev VV, Kekuda R, Pena CBA;
XX  Zhong M, Ganglilil BA, Taupier RJ;
XX  WPI, 2002-713508/77.
XX  P-PSDB; ABG97498.
XX
XX  New NOVX polypeptides and polynucleotides, useful for preventing,
XX  diagnosing or treating NOVX-associated disorders, e.g. diabetes, multiple
XX  sclerosis, atherosclerosis, cancer, infections, osteoporosis or
XX  Parkinson's disease.
XX
XX  Claim 22; Page 139; 266pp; English.
XX
XX  The present invention relates to a new polypeptide (NOVX). The NOVX
XX  polypeptide, nucleic acid and antibody are useful in the manufacture of a
XX  medicament for treating a syndrome associated with a human disease,
XX  preferably a NOVX-associated disorder. The NOVX nucleic acids,
XX  polypeptides and antibodies are useful for treating, preventing or
XX  diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
XX  syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
XX  hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
XX  palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-
XX  telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
XX  obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
XX  infertility, inflammatory bowel disease, atherosclerosis, hypertension,
XX  scleroderma, haemophilia, diabetes, pancreatitis, autoimmune disease,
XX  acthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
XX  infections, or graft-versus-host disease. The nucleic acids and
XX  polypeptides may also be used as targets for the identification of small
XX  molecules that modulate or inhibit e.g. neurogenesis, cell
XX  differentiation, cell proliferation, haematopoiesis, wound healing and
XX  angiogenesis, in gene therapy, in generation of antibodies that bind
XX  immunospecifically to NOVX substances for use in therapeutic or
XX  diagnostic methods. The nucleic acids are further used as hybridisation
XX  probes, in chromosome mapping, tissue typing, preventive medicine, and
XX  pharmacogenomics. The present nucleic acid sequence encodes a human NOVX
XX  protein of the invention
XX
XX  Sequence 619 BP; 97 A; 219 C; 207 G; 96 T; 0 U; 0 Other;

```


Alignment Scores:

Pred. No.: 9.41e-58 Length: 645
 Score: 593.00 Matches: 106
 Percent Similarity: 79.14% Conservative: 42
 Best Local Similarity: 56.68% Mismatches: 39
 Query Match: 61.64% Indels: 0
 DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x ABS78741 (1-645)

QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeuGly 21
 DB 23 GCCATGACACCTTCACAGCAAGAGCTGCTCTGACGGCGCAAGAAAGACTCTCAGT 82
 QY 22 LysMetValSerLysSerLleAlaThrThreuleleAspAspThrSerSerGluValleu 41
 DB 83 AAGATGGCGTTCACAGGAGGTGGCGCTGCTGGATGACACAGAGTGAAGTGTCTG 142
 QY 42 AspGluLeuTyrArgValThrArgLutyrThrGlnAsnLysLysGluAlaGluLysLys 61
 DB 143 GATAGCTGTACCGCCGACACGAGGAGTTCACGGCAGCCGCAAGAGGCGCCAGAGATG 202
 QY 62 IleLysAsnLeuLleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPhe 81
 DB 203 CTCAGAACTGCTGCAAGAGTGGCCCTGAGAGCTGGAGACTGCTGCTGGGGGACCACTG 262
 QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
 DB 263 GCGCGTGAAGAGTGGCGCTGCTGCGGCTGCCACCGAGGCGCGCTGCTGCGCATG 322
 QY 102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
 DB 323 ACGGCGCTGACCTTCACAGTGGACTTCACTTCAACCGCGCGCTGCTGCGCGCGG 382
 QY 122 LeuAsnGluCysArgLysLysLeuHisGlnIleGlnArgHisLeuThrAlaLysSer 141
 DB 383 CTGCTCGAGTGGCCGACCTGCTGACAGGCGCTGGGTCCCGACCTGACCGCCAGATCC 442
 QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
 DB 443 CACGGCGGATCAACACGCTTGGCCACCTGACGACCTGCGACTTCTGCGCGCTC 502
 QY 162 TyrAsnProPheGlyAsnPheLysPheHisLeuGlnLysLeuCysAspGlyLleAsnLys 181
 DB 503 TACGGCGCCCGCCAGACCTTACCGCTCCACCTGCGAGATCTGCGAGGCGCTGCGCGG 562
 QY 182 MetLeuAspGluGluAsnIle 188
 DB 563 ATGCTGACGAGGCGACCTC 583

RESULT 22
 AAL39626 standard; DNA; 2108 BP.
 XX
 AAL39626;
 XX
 DT 05-SEP-2002 (first entry)
 XX
 DE Human secreted protein DNA SEQ ID NO 70.
 XX
 KW Antiartherosclerotic; cyostatic; HIV; antiallergic; antianaemic;
 KW antidiabetic; cardiac; vasotonic; neuroprotective; nootropic; SECP;
 KW anticonvulsant; antiparkinsonian; cerebroprotective; antiinflammatory;
 KW immunosuppressive; human secreted protein; cell proliferative disorder;
 KW arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;
 KW allergy; anaemia; asthma; cardiovascular disease; developmental disorder;
 KW ischaemic heart disease; congestive heart failure; neurological disorder;
 KW renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;
 KW Parkinson's disease; epilepsy; stroke; knockin humanised animal;
 KW transgenic animal; gene therapy; gene; ds.
 XX
 OS Homo sapiens.

XX
 FN W0200238602-A2.
 XX
 PD 16-MAY-2002.
 XX
 PF 08-NOV-2001; 2001WO-US047420.
 XX
 XX 08-NOV-2000; 2000US-0247505P.
 PR 09-NOV-2000; 2000US-0248642P.
 PR 16-NOV-2000; 2000US-0249824P.
 PR 21-NOV-2000; 2000US-0252824P.
 PR 08-DEC-2000; 2000US-0254305P.
 PR 18-DEC-2000; 2000US-0256448P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PI Yue H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Walla NK;
 PI Sanjwalwa M, Thornton M, Elliott VS, Lu Y, Gierzen KJ, Burford N;
 PI Ding L, Hafalia AA, Tang YT, Bandman O, Warren BA, Honchell CD;
 PI Lu DM, Thangavelu K, Lee S, Xu Y, Yang J, Lai PG, Tran B;
 PI Ison CH, Duggan BM, Saperstein SK;
 XX
 DR WPI: 2002-519236/55.
 DR P-PSDB; AAO21665.
 XX
 PT Human secreted proteins and polynucleotides for diagnosing, treating or
 PT preventing disorders of cell proliferative, cardiovascular,
 PT developmental, neurological and autoimmune/inflammatory disorders.
 XX
 PS Claim 5; Page 195-196; 229pp; English.

CC The invention relates to an isolated human secreted protein (SECP)
 CC polypeptide from 63 fully defined protein sequences given in the
 CC specification. The polypeptide is useful for the diagnosing/treating of a
 CC disease with decreased/overexpression of SECP. Examples of disorders
 CC associated with abnormal expression of SECP include a cell proliferative
 CC disorder e.g. arteriosclerosis, cancers; autoimmune/inflammatory
 CC disorder, AIDS, allergies, anaemia, asthma; cardiovascular disease e.g.
 CC congestive heart failure, ischaemic heart disease; developmental disorder
 CC e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.
 CC Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.
 CC The SECP polynucleotide and polypeptide are further useful for analysing
 CC the proteome of a tissue or a cell type. The polynucleotide is useful for
 CC creating knockin humanised animals (pigs) or transgenic animals (mice or
 CC rats) to model human disease, and for somatic or germ-line gene therapy,
 CC and further for generating hybridisation probes useful in mapping the
 CC naturally occurring genomic sequence. This polynucleotide sequence
 CC represents the DNA of a human secreted protein of the invention
 XX
 XX Sequence 2108 BP; 408 A; 602 C; 636 G; 462 T; 0 U; 0 Other;
 SQ

Alignment Scores:
 Pred. No.: 4.47e-57 Length: 2108
 Score: 593.00 Matches: 106
 Percent Similarity: 79.14% Conservative: 42
 Best Local Similarity: 56.68% Mismatches: 39
 Query Match: 61.64% Indels: 0
 DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x AAL39626 (1-2108)

QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLeuGly 21
 DB 427 GCCATGACACCTTCACAGCAAGAGCTGCTCTGACGGCGCAAGAAAGACTCTCAGT 486
 QY 22 LysMetValSerLysSerLleAlaThrThreuleleAspAspThrSerSerGluValleu 41
 DB 487 AAGATGGCGTTCACAGGAGGTGGCGCTGCTGGATGACACAGAGTGAAGTGTCTG 546
 QY 42 AspGluLeuTyrArgValThrArgLutyrThrGlnAsnLysLysGluAlaGluLysLys 61
 DB 547 GATAGCTGTACCGCCGACACGAGGAGTTCACGGCAGCCGCAAGAGGCGCCAGAGATG 606

QY 162 TyrAsnProPheGlyAsnPhelyProHisIeuGlnIysLeuCyAspGlyIleAsnIys 181
DB 593 TACGGCCCCCGGACCCCTACCGCTCCCGCAGATCTGCAGAGCCCTGCGCCG 652
QY 182 MetLeuAspGluGluAsnIle 188
DB 653 ATGCTGACGAGGCGACGCTC 673
RESULT 24
ADQ86751
ID ADQ86751 standard; cDNA; 2186 BP.
XX
AC ADQ86751;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #3626.
XX
KM human; tumour-associated antigenic target; TAT; cytosolic; gene therapy;
KM cancer; cell proliferative disorder; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2004060270-A2.
XX
PD 22-JUL-2004.
XX
PF 15-OCT-2003; 2003WO-US029126.
XX
PR 18-OCT-2002; 2002US-0418988P.
PA (GETH) GENENTECH INC.
PA (WOTD/) WU T D.
PA (ZHOU/) ZHOU Y.
XX
PI Wu TD, Zhou Y;
XX
DR WPI; 2004-534300/51.
XX
PT New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
PS Claim 1; SEQ ID NO 3626; 5504bp; English.
XX
CC The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a) - (c); or (e) a sequence that hybridises to (a) -
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a

CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 4.69e-57 Length: 2186
Score: 593.00 Matches: 106
Percent Similarity: 79.14% Conservative: 42
Best Local Similarity: 56.68% Mismatches: 39
Query Match: 61.64% Indels: 0
DB: 12 Gaps: 0
US-10-627-571-2 (1-188) x ADQ86751 (1-2186)
QY 2 AlaThrAspValPheAsnSerIysAsnLeuAlaValGlnIleGlnIysIleLeuGly 21
DB 113 GCCATGGACACCTTCAGCACCAAGCCTGCTCTCCAGCGGAGAAAGCTCTGAGT 172
QY 22 LysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
DB 173 AAGATGGCGCTCCAAAGCAGGACGTGGCGCTGCTGTGATGACACCACTGATGCTGCTG 232
QY 42 AspGluLeuValThrValThrValThrValThrValThrValThrValThrValThrValThr 61
DB 233 GATGACTGTAACCGCCGACCAAGGAGATTCAACCGACCGGACGAGAGGCCGAGAGATG 292
QY 62 IleIysAsnLeuIleIysThrValIleIysLeuAlaIleLeuValThrValThrValThrVal 81
DB 293 CTCAGAAACCTGTCAAGGTGGCGCTGAAAGCTGGAGCTGTGTGTGGGACCAAGCTG 352
QY 82 AsnGluAspGluLeuAlaLeuMetGluIysPheValIysValIleGlnLeuAlaMet 101
DB 353 GCGGCTGAGGAGCTGGCGCTGCTGCGCGGCTTCCGACCGGCGGCTGCGCGG 412
QY 102 ThrValValSerPheHisGlnValAspValThrPheAspArgAsnValLeuSerArgLeu 121
DB 413 ACGGCGGTGAGCTTCCACCAAGGTGACTTCACTTGCACCGGCGGCTGCGCGG 472
QY 122 LeuAsnGluCyAspArgLeuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIysSer 141
DB 473 CTGCTGAGTGCAGCGACCTGTGACCAAGCCGTGGGTCCCACTGACCGGCAAGTCC 532
QY 142 HisGluArgValAsnAsnValPheAspHisPheSerAspCyAspGluPheLeuAlaLeu 161
DB 533 CAGGCGGCTATCAACACGCTGTTCGCGCACCTAGCCGACTTGCACCTTCTGCGGCTC 592
QY 162 TyrAsnProPheGlyAsnPhelyProHisIeuGlnIysLeuCyAspGlyIleAsnIys 181
DB 593 TACGGCCCCCGGACCCCTACCGCTCCCGCAGATCTGCAGAGCCCTGCGCCG 652
QY 182 MetLeuAspGluGluAsnIle 188
DB 653 ATGCTGACGAGGCGACGCTC 673
RESULT 25
ADQ84271
ID ADQ84271 standard; cDNA; 2186 BP.
XX
AC ADQ84271;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1085.

KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
KM cancer; cell proliferative disorder; gene; ss.
XX Homo sapiens.
XX WO2004060270-A2.
XX
XX 22-JUL-2004.
XX
XX 15-OCT-2003; 2003WO-US029126.
XX
XX 18-OCT-2002; 2002US-0418988P.
XX
XX (GENTH) GENENTECH INC.
XX (WUTD/) WU T D.
XX (ZHOU/) ZHOU Y.
XX
XX Wu TD, Zhou Y;
XX
XX MPI; 2004-534300/51.
XX
XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
PT preventing or treating cell proliferative disorders such as cancer.
XX
XX Claim 1; SEQ ID NO 1085; 5504pp; English.
XX
XX The present invention describes an isolated tumour-associated antigenic
CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
CC sequences (see SEQ ID NO1 to 4622); (b) the full-length coding region of
CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
CC sequence identity to (a) -(c); or (e) a sequence that hybridises to (a) -
CC (c). Also described: (1) an expression vector comprising the above
CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
CC a process for producing a polypeptide; (4) an isolated polypeptide
CC comprising: (a) an amino acid sequence encoded by any of the above
CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
CC length coding region of the above nucleotide sequences; or (c) a sequence
CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
CC an isolated antibody that binds to the above polypeptide; (7) a process
CC for producing the antibody; (8) an isolated oligopeptide that binds to
CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
CC binding organic molecule that binds to the above polypeptide; (10) a
CC composition of matter comprising the above (chimeric) polypeptide,
CC antibody, oligopeptide or TAT binding organic molecule, in combination
CC with a carrier; (11) an article of manufacture comprising a container and
CC the composition of matter contained within the container; (12) methods of
CC inhibiting the growth of a cell that expresses the above protein, where
CC the growth of the cell is at least in part dependent upon a growth
CC potentiating effect of the above protein; (13) a method of
CC therapeutically treating a mammal having a cancerous tumour comprising
CC cells that express the above protein; (14) a method of determining the
CC presence of a protein in a sample suspected of containing the protein
CC described above; (15) methods of diagnosing the presence of a tumour in a
CC mammal; (16) a method for treating or preventing a cell proliferative
CC disorder associated with increased expression or activity of the above
CC protein; and (17) a method of binding an antibody, oligopeptide or
CC organic molecule to a cell that expresses the protein described above.
CC The TAT sequences have cytostatic activities, and can be used in gene
CC therapy. The composition and methods are useful for diagnosing,
CC preventing or treating cancer. The composition is also used for preparing
CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.
XX
SQ Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.69e-57 Length: 2186
Score: 593.00 Matches: 106
Percent Similarity: 79.14% Conservative: 42
Best Local Similarity: 56.68% Mismatches: 39
Query Match: 61.64% Indels: 0

DB: 13 Gaps: 0
US-10-627-571-2 (1-188) x ADQ84271 (1-2186)
QY 2 AATTAAAPVAlPheAaMSerLYsAsnLeuAlaValGlnAlaGlnLYsLeuIleuGly 21
DB 113 GCCATGACACCTTGACGACCAAGAGCCTGGCTGACAGCCGACAGAGAGCTCTGAGT 172
QY 22 LysMetValSerLYsSerLLeAlaThrThrLeuLLeAaPAPThSerSerGluValLeu 41
DB 173 AAGATGGCGTCCAGAGCGAGTGGCGCTGGTGGATGACACGACGATGAGTGCTG 232
QY 42 AaPGLuLeuLYrArGVAlThzArGVuLYrThrGlnAaMlyLYsGluAlaGluLYsLYs 61
DB 233 GATGAGCTGTACCGGCCACGAGGAGTTCACAGCCGACGCCAAGAGGCCGACAAGATG 292
QY 62 lIeLYsAsnLeuIlleLYsThrValIlleLYsLeuAlaIlleLYrArGVsAsnGlnPhe 81
DB 293 CTCAGAAACCTGGTCAAGGTGGCCCTGAAGCTGGACTGCTGCGGGAGACGACTG 352
QY 82 AaGlnAaPGLuLeuAlaLeuMetGluLYsPheLYsLYsValHisGlnLeuAlaMet 101
DB 353 GCGCGTGAAGAGCTGGCGCTGGCGGCTTCGCCACGCCGCCGCTGCTGGCATG 412
QY 102 ThrValValSerPheHisGlnValAspTYrThrPheAsPArGVsValLeuSerArgLeu 121
DB 413 ACGGCGGTGACGTTCCACAGGTGACCTTCACCTTCGACCGCGGTGCGCGCGG 472
QY 122 LeuAsnGluLYrArGVAlMetLeuHisGlnIlleGlnAaGHisLeuThrAlaLYsSer 141
DB 473 CTGCTCGAGTGGCGGACCTGTCGACCAAGCGCGGTGCCACCTGACCGCCAAAGTCC 532
QY 142 HisGlyArGVAlAsnAsnValPheAspHisPheSerAaPYrGluPheLeuAlaLeu 161
DB 533 CACGGCGGATACACACGATGTCGACCACTGACCGATGCGACTTCCTGGCTGGCTC 592
QY 162 TYrAsnProPheGlyAsnPheLYsProHisLeuGlnLYsLeuCYsAspGlylLeAsnLYs 181
DB 593 TACGGCGCGCGGAGCCCTACCGCTCCACCTGCGCAGAGATGTGGAGGCGCTGGGCGG 652
QY 182 MetLeuAspGluGluAsnIle 188
DB 653 ATGCTGGACGAGGCGACCTC 673

RESULT 26
AAH07641
ID AAH07641 standard; cDNA; 587 BP.
XX
XX AAH07641;
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA clone (5'-primer) SEQ ID NO:4476.
XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX Homo sapiens.
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99UP-00248036.
XX 27-AUG-1999; 99UP-00300253.
XX 11-JAN-2000; 2000JP-00118776.
XX 02-MAY-2000; 2000JP-00183767.
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Iehi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 1; SEQ ID NO 4476; 2537bp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification, where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX

XX Sequence 587 BP; 178 A; 146 C; 140 G; 116 T; 0 U; 7 Other;

Alignment Scores:

Pred. No.:	3,99e-57	Length:	587
Score:	587.00	Matches:	123
Percent Similarity:	93.94%	Conservative:	1
Best Local Similarity:	93.18%	Mismatches:	7
Query Match:	61.02%	Indels:	1
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x AAH07641 (1-587)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLileu 20
 DB 140 ATGGCCACAGATGCTTTAAATCCAAACCTGCGCTCAGGACAAAGAAATCTTG 199
 QY 21 GILySeMetValSerLysSerLilealathThrLenuleAspAspThrSerSerGluVal 40
 DB 200 GGRNAAATGCTGTCAAATCCATCGCCACCACTTAATACACACACAAATATGAGTG 259
 QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLys 60
 DB 260 CTGGAATGAGCTCTTACAGAGTACAGGAGCTACACCAAAACAAAGAGGAGGAGAG 319
 QY 61 LysILlLysAsnLenuleLysThrValLleLysLeuAlaLleLenuTyrArgAsnAnGln 80
 DB 320 ATATCAAGAACCTCATCAAGACAGATCATCAAGCTGCGCTTTATAGAAATATCAG 379
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 DB 380 TTTAACTCAAGATGAGTATGATGAGGAAATTTAAGAAAGAAAGTTATCAGCTTGCT 439
 QY 101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 DB 440 ATACCCGTGNCAGTTTCCAAAGAGGGGTTTAACTTTGACCGAATGTGTATCCAG 499
 QY 121 LeuLeuAsnGluCys-ArgGluMetLeuHisGln 131
 DB 500 CTGTTAAATGAATGCANAGAAATGCTGCACCAA 533

RESULT 27
 ABK34943
 ID ABK34943 standard; cDNA; 1113 BP.

AC ABK34943;

DT 08-MAY-2002 (first entry)

DE Human cDNA encoding secreted protein #81.

KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
 KW viral infection; bacterial infection; fungal infection; diabetes; achma;
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
 KW tissue regeneration; wound healing; burn; haematopoiesis;
 KW myeloid cell deficiency; lymphoid cell deficiency.

OS Homo sapiens.

PN MO200177288-A2.

PD 18-OCT-2001.

PF 29-MAR-2001; 2001WO-US010224.

PR 06-APR-2000; 2000US-0195582P.

PA (GENY) GENNETICS INST INC.

PI Wong Gg, Clark HF, Fecttel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulukota K, Graham JR;

PI WPI; 2002-179321/23.

XX Five hundred and ninety two polynucleotides derived from a variety of
 XX human tissue sources which encode secreted proteins, useful for treating
 XX immune deficiencies and disorders such as autoimmune disorders.

PS Claim 1; Page 109; 372pp; English.

CC The invention relates to 592 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins. The polynucleotides can be used as probes for the
 CC identification and isolation of full length cDNA and genomic DNA. The
 CC polynucleotides and proteins can also be used as nutritional supplements.
 CC The proteins are useful in the treatment of various immune deficiencies
 CC and disorders such as viral infections, bacterial infections, fungal
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
 CC and conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment of
 CC burns, incisions and ulcers. The proteins are also useful for regulating
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
 CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention

XX Sequence 1113 BP; 282 A; 311 C; 304 G; 216 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	5,97e-50	Length:	1113
Score:	527.00 <td>Matches:</td> <td>99 </td>	Matches:	99
Percent Similarity:	77.22% <td>Conservative:</td> <td>40 </td>	Conservative:	40
Best Local Similarity:	55.00% <td>Mismatches:</td> <td>39 </td>	Mismatches:	39
Query Match:	54.78% <td>Indels:</td> <td>2 </td>	Indels:	2
DB:	6	Gaps:	1

US-10-627-571-2 (1-188) x ABK34943 (1-1113)

```

QY 6 PheAsnSerIlyAsnLeuAlaValGlnAlaGlnIlyLeuGlnIlySer 25
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 72 TTCAGCTCAAAAGAGCTGGCACTGCAGACAGAAAGACTGATGATGGCGGT 131
QY 26 LysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTy 45
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 132 CGCTGTGGCTCATCTCTTCATAGATGAGACAAAGCATGAGGTGCTAGATGAGCTTAC 191
QY 46 ArgValThrArgGlnIlyThrGlnAsnIlyLysGlnAlaGlnIlyLysValLeu 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 192 CGGTGTCCAAAGAGTACAGCAGCAGCGGCCCGCCCGCATGATCAAGAGCTTG 251
QY 66 IleTyThrValIleLysLeuAlaIleLeuTyArgAsnAsnGlnPheAsnGlnAspGlu 85
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 252 ATCAAAAGTGGCCATCAAGGTGGCTGTGCTGCACCGCAATGGCTTTGGCCCACTGAG 311
QY 86 LeuAlaLeuMetGlnLysPheLysIlyValIleGlnLeuAlaMetThrValValSer 105
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 312 CTGGCCCTGGCTACCGGCTTTGGCCAGAGCTGCGCAGGGGTCCATGACGGCACTTAC 371
QY 106 PheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCy 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 372 TTGTGTAGGTACCTTCACCTTCGAGGCTGCTGTGCTGCTGCTGACCGAGTGC 431
QY 126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgVal 145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 432 CGGAGTGTCTGCTTAAGTTGGTGGAAACACACCTTACCGCCCAAGTCAATGCGCCATC 491
QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTyArgProPhe 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 492 CGCACAGTGTGTATCATCTCTCTGACCCAGAGTCTGCTCAGCGCCCTATGGGCT--- 548
QY 166 GlyAsnPheLysProHisLeuGlnIlyLysLeuCysAspGlyIleAsnIlyMetLeuAspGlu 185
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 549 ---GACTTACTCAGCACCCTTGGCAAGATCTGTGACGSACTCAGAAAGCTGTTAGCGAA 605

RESULT 28
AAD38699
ID AAD38699 standard; cDNA; 1165 BP.
XX
AC AAD38699;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human LP237 secreted protein encoding cDNA.
XX
KW Human; secreted protein; atherosclerosis; Alzheimer's disease; LP237;
KW diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
KW rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
KW reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
KW gene therapy; neoplasm; transgenic; psoriasis; ischemia; carcinoma;
KW gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 100..654
FT /tag= a
FT /product= "human LP237 secreted protein"
FT sig_peptide 100..180
FT /*tag= b
FT mat_peptide 181..651
FT /*tag= c
FT /product= "mature human LP237 secreted protein"
XX
PN WO200226801-A2.
XX
PD 04-APR-2002.
XX
PF 14-SEP-2001; 2001WO-US026026.
XX
PR 28-SEP-2000; 2000US-023608AP.
XX

```

```

PA (ELT ) LILLY & CO ELI.
XX
XX Su EW, Wang H;
XX
DR WPI: 2002-471259/50.
XX
P-PSDB; AAE23984.
XX
PT Novel proteins and polynucleotides of secreted proteins useful for
PT treating various diseases e.g. rheumatoid arthritis, cancer, psoriasis,
PT diabetic retinopathy, arteriosclerosis, ischemia or reperfusion injury.
XX
PS Claim 1; Page 140-141; 145pp; English.
XX
CC The invention relates to human secreted polypeptides designated LP095,
CC LP191, LP217, LP220, LP221, LP222, LP225, LP237 or LP238 and nucleic acid
CC molecules encoding such polypeptides. Novel secreted proteins of the
CC invention are used for treating diseases such as atherosclerosis,
CC Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis,
CC arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe
CC combined immunodeficiency, ischaemia, carcinomas, haemolytic anaemia,
CC reperfusion injury, neoplasms and cancer especially liver cancer. They
CC are also used for wound healing. Polynucleotides of the invention can be
CC used to generate transgenic animals or knock out animals, which in turn,
CC are useful in the development and screening of therapeutically useful
CC reagents for use in the treatment of diseases associated with LP
CC polypeptide associated activity. They are also used in gene therapy. The
CC present sequence is human LP237 secreted protein encoding cDNA
XX
SQ Sequence 1165 BP; 305 A; 319 C; 312 G; 229 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 6,34e-50 Length: 1165
Score: 527.00 Matches: 99
Percent Similarity: 77.22% Conservative: 40
Best Local Similarity: 55.00% Mismatches: 39
Query Match: 54.78% Indels: 2
DB: Gaps: 1
XX
US-10-627-571-2 (1-188) x AAD38699 (1-1165)
QY 6 PheAsnSerIlyAsnLeuAlaValGlnAlaGlnIlyLeuGlnIlySer 25
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 109 TTCAGCTCAAAAGAGCTGGCACTGCAGACAGAAAGACTGATGATGGCGGT 168
QY 26 LysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTy 45
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 169 CCCTGTGGCTCATCTCTTCATAGATGAGACAAAGCATGAGGTGCTAGATGAGCTTAC 228
QY 46 ArgValThrArgGlnIlyThrGlnAsnIlyLysGlnAlaGlnIlyLysValLeu 65
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 229 CGGTGTCCAAAGAGTACAGCAGCAGCGGCCCGCCCGCATGATCAAGAGCTTG 288
QY 66 IleTyThrValIleLysLeuAlaIleLeuTyArgAsnAsnGlnPheAsnGlnAspGlu 85
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 289 ATCAAAAGTGGCCATCAAGGTGGCTGTGCTGCACCGCAATGGCTTTGGCCCACTGAG 348
QY 86 LeuAlaLeuMetGlnLysPheLysIlyValIleGlnLeuAlaMetThrValValSer 105
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 349 CTGGCCCTGGCTACCGGCTTTGGCCAGAGCTGCGCAGGGGTCCATGACGGCACTTAC 408
QY 106 PheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCy 125
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 409 TTGTGTAGGTACCTTCACCTTCGAGGCTGTGCTGCTGCTGCTGACCGAGTGC 468
QY 126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgVal 145
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 469 CGGAGTGTCTGCTTAAGTTGGTGGAAACACACCTTACCGCCCAAGTCAATGCGCGCATC 528
QY 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTyArgProPhe 165
   |||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 529 CGCACAGTGTGTATCATCTCTGACCCAGAGTCTCAGCGCCCTCTATGGGCT--- 585
QY 166 GlyAsnPheLysProHisLeuGlnIlyLysLeuCysAspGlyIleAsnIlyMetLeuAspGlu 185

```

Db 586 ---GACCTTCACTGACGACCTTGGCAAGATCTGTGACGACTCAGAGAGCTGCTTACGAA 642

RESULT 29

AB211655

ID AB211655 standard; cDNA, 1175 BP.

XX

AC AB211655;

XX

DT 20-JAN-2003 (first entry)

XX

DE Human polynucleotide SEQ ID NO 537.

XX

KW Human, genome mapping, gene therapy, food supplement, virus, fungus;

KW cell-proliferative disorder; neurodegenerative disease; bacteriophage;

KW Parkinson's disease; Alzheimer's disease; autoimmune disease;

KW multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;

KW arthritis; cytostatic; immunomodulator; neurotropic; neuroprotective;

KW antiparkinsonian; antidiabetic; immunosuppressive; dermatological;

KW hemostatic; vulnereary; fungicide; antibacterial; virucide; protozoacide;

KW antithalitic; gene; ss.

XX

OS Homo sapiens.

XX

PN WO200270539-A2.

XX

PD 12-SEP-2002.

XX

PF 05-MAR-2002; 2002WO-US005095.

XX

PR 05-MAR-2001; 2001US-00799451.

XX

PA (HYSB-) HYSBQ INC.

XX

PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;

PI Wehrman T, Wang J, Wang D, Drmanac RT;

DR WPI; 2002-759812/82.

DR P-PSDB; ABP69438.

XX

PT New polynucleotides comprising sequences assembled from expressed

PT sequence tags (ESTs), useful for treating cell-proliferative,

PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet

PT or coagulation disorders.

XX

PS Claim 1; SEQ ID NO 537; 1012bp + Sequence Listing; English.

XX

CC The invention relates to an isolated polynucleotide (I) comprising a

CC nucleotide sequence selected from any of 948 sequences (AB211119-

CC AB212066) or their mature protein coding portion, active domain coding

CC protein or complementary sequences. The polynucleotides are useful for

CC identifying expressed genes or for physical mapping of human genome. The

CC encoded polypeptides (ABP68902-ABP68943) are useful as molecular weight

CC markers, as a food supplement, for generating antibodies, in medical

CC imaging, screening and diagnostic assays and for treating cell-

CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's

CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,

CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,

CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver

CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),

CC arthritis, etc. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SO Sequence 1175 BP; 306 A; 324 C; 313 G; 232 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 6,41e-50 Length: 1175

Score: 527.00 Matches: 99

Percent Similarity: 77.22% Conservative: 40

Best Local Similarity: 55.00% Mismatches: 39

Query Match: 54.78% Indels: 2

DB: 6 Gaps: 1

US-10-627-571-2 (1-188) x AB211655 (1-1175)

Qy 6 PheAsnSerLyAsnLeuAlaValGlnAlaGlnLySerIleLeuGlyLyMetValSer 25

122 TTCAGCTCAAAAGCCTGGCACTGCACAGAGAAAGACTTACTAGTAAGATGGCGGT 181

Qy 26 LysSerIleAlaThrThrIleuIleAspAspThrSerSerGluValLeuAspGluLeuTyr 45

182 CGCTGTGTGCTCATCTCTTATAGATGACACAAAGCATGAGTGTCTAGATGAGCTTAC 241

Qy 46 ArgValThrArgGluTyrThrGlnAsnLysGluAlaGluLysLysIleLysAsnLeu 65

242 CGTGTGTCCAAAGAGATACACGACACCGCCGCCAGCCGCGCTGATCAAGACACTG 301

Qy 66 IleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnIleAspGlu 85

302 ATCAAAAGTGGCAATCAAGAGGCTGTGTGACCGGCAATGCTCTTGGCCCCAGTTAG 361

Qy 86 LeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAlaMetThrValIleSer 105

362 CTGGCCCTGGCTACCCGCTTTCGCCAAGAGCTGGGCAAGGCTCCATGACGGACTTAC 421

Qy 106 PheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125

422 TTGTGTGAGGTAGACCTTCACTTCAGAGCTGTGTCTGTCTGTGCTGTGACCGAGTGC 481

Qy 126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgVal 145

482 CGGAGTGTGTGTGTAGAGTGTGTGTGAACACACCTTACCGCCCAAGTCACTGACCGCATC 541

Qy 146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleLeuTyrAsnProPhe 165

542 CGCCAGCTGTGTATATCTTCTGTGACCAAGCTCTCTACGCGCTCTATGGCGCT--- 598

Qy 166 GlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeuAspGlu 185

599 ---GACCTCACTCAGCAGCTTGGCAAGATCTGTGACGAGACTCAGAAAGCTGTAAGACGAA 655

RESULT 30

ID ADM44173 standard; cDNA, 1175 BP.

XX

AC ADM44173;

XX

DT 03-JUN-2004 (first entry)

XX

DE Novel human arginine-rich protein cDNA #537.

XX

KW ss; gene; human; arginine-rich protein; cancer; inflammation;

KW genetic disorder.

XX

OS Homo sapiens.

XX

PN US2004053250-A1.

XX

PD 18-MAR-2004.

XX

PF 21-NOV-2002; 2002US-00302172.

XX

PR 05-MAR-2001; 2001US-00799451.

PR 05-MAR-2002; 2002WO-US005095.

PR 20-AUG-2002; 2002US-00225251.

XX

PA (TANG/) TANG Y T.

PA (XUEA/) XUE A.

PA (DRMA/) DRMANAC R T.

PI Tang YT, Xue A, Drmanac RT;

DR WPI; 2004-238579/22.


```

Db      172 CGCTCTGCTATCTCTTCATGAGACAAAGCAAGTGAAGCTGATGAGCTTAC 231
      46 ArgValThrArgIuYrThrGlnAsnLysGluAgluLysLysIleLysAsnLeu 65
      232 CGTGTGTCAAGAGGTACACGACACGCGGCCAGGCCGCGATCAAGACCTG 291
Qy      66 IleLysThrValIleLysLeuAlaIleLeuYrArgAsnAsnGlnPheAsnGlnAspGlu 85
      292 ATCAAGGTGCGCATCAAGGTGGCTGTGCTGCAACGCAATGCTCTTGGCCCAAGTAG 351
Qy      86 LeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMetThrValSer 105
      352 CTGGCCTTGGCTACCCCTTTCGACAGAGCTGGCGAGGGTGCATGAGCGACCTTAC 411
Qy      106 PheHisGlnValAspYrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125
      412 TTGTGTAGGTAGACCTTCACTTCGAGGCTGCTGCTTGGCTGGCTGACCGAGTGC 471
Qy      126 ArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSerHisGlyArgVal 145
      472 CGGATGTGCTGTAGAGTTGTGTGAACACCACTCAAGCCCAAGTCAATGGCGCATC 531
Qy      146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeuYrAsnProPhe 165
      532 CGCCACGTGTTGATTCATCTCTCTGACCCAGCTGTCTCAAGGCTCTTATGGGCT-- 588
Db      166 GlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLysMetLeuAspGlu 185
      589 ---GACTTCACTACGACCTTTGGCAATCTGTGACCGACTCAGAAAGCTCTTAGACGA 645
RESULT 32
ABLO7183
ID      ABL07183 standard; cDNA; 829 BP.
XX
AC      ABL07183;
XX
DT      26-MAR-2002 (first entry)
XX
DE      Drosophila melanogaster expressed polynucleotide SEQ ID NO 16031.
XX
KW      Drosophila; developmental biology; cell signalling; insecticide;
XX
KW      pharmaceutical; gene; ss.
XX
OS      Drosophila melanogaster.
XX
PN      WO200171042-A2.
XX
PD      27-SEP-2001.
XX
PF      23-MAR-2001; 2001WO-US009231.
XX
PR      23-MAR-2000; 2000US-0191637P.
XX
PR      11-JUL-2000; 2000US-00614150.
XX
PA      (PEKE ) PE CORP NY.
XX
PI      Venter JC, Adams M, Li PWD, Myere EW,
XX
DR      WPI; 2001-656860/75.
XX
DR      P-PDB; ABB63080.
XX
PT      New isolated nucleic acid detection reagent for detecting 1000 or more
XX
PT      genes from Drosophila and for elucidating cell signalling and cell-cell
XX
PT      interactions.
XX
PS      Claim 1; SEQ ID NO 16031; 21PP + Sequence Listing; English.
XX
CC      The invention relates to an isolated nucleic acid detection reagent
XX
CC      capable of detecting 1000 or more genes from Drosophila. The invention is
XX
CC      useful in developmental biology and in elucidating cell signalling and
XX
CC      cell-cell interactions in higher eukaryotes for the development of
XX
CC      insecticides, therapeutics and pharmaceutical drugs. The invention

```

```

CC      discloses genomic DNA sequences (ABU16176-ABU30511), expressed DNA
CC      sequences (ABLO1840-ABLO16175) and the encoded proteins (ABB57737-
CC      ABB72072). The sequence data for this patent did not form part of the
CC      printed specification, but was obtained in electronic format directly
CC      from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 829 BP; 229 A; 208 C; 210 G; 182 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.:      8,09e-36      Length:      829
Score:      401.00      Matches:      79
Percent Similarity:      65.43%      Conservative:      44
Best Local Similarity:      42.02%      Mismatches:      65
Query Match:      41.68%      Indels:      0
DB:      4      Gaps:      0
US-10-627-571-2 (1-188) x ABL07183 (1-829)
Qy      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAgluLysLysIleLeu 20
      263 ATGGCGAGCAATGCTTCAAGTGCACGATATGGGCTGGCGCGAGAAAGATCTTC 322
Db      21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspHisSerSerGluVal 40
      323 TCAGCATGGCCACAAATAATCGCAAGACCTTCATCATGATGCAAGCGGCTGCTG 382
Qy      41 LeuAspGluLeuYrArgValThrArgIuYrThrGlnAsnLysGluAgluLys 60
      383 CTGACACACCTCTACAGGCTCTGCAGATGCACACGGGGAACAGGCCAAGCGCAAGG 442
Db      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuYrArgAsnGln 80
      443 CTGATCAAGAACATCATCAAGATTGTGATCAAGATGCTGTCCACCGCAACATCAG 502
Qy      81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
      503 TTCAGGACGAGAGGTGCAAGAAAGCGGAGCTCTTTAAGAAAGTTTCAAAACACGCGAG 562
Db      101 MetThrValValSerPheHisGlnValAspYrThrPheAspArgAsnValLeuSerArg 120
      563 CTGTCCATCATATCATCTTACAGAGTGCATTCACCTTGCACCTGCCGTRACCTCAAAAG 622
Qy      121 LeuLeuAsnGluCysArgLysLeuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      623 TCATATAGCCGAATCCGAAAGTGGCGCTGAAGTGCATTTGTCAGCGCATCTCACAGAAAG 682
Db      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
      683 TCCTGTGGCCGATCGACGAGGTGTCATTTCTTGGCGAGAGGCGCTGTGAGAGCG 742
Qy      161 LeuYrAsnProPheGlnAsnPheLysPheHisLeuGlnLysLeuCysAspGlyIleAsn 180
      743 GCTTTCGACCAAGATTCCTCCACCGAGAGTGTGAGAAAGTTGTGGCGGACATTAAC 802
Db      181 LysMetLeuAspGluGluAsnIle 188
      803 GCAGCATGGAAGACGGAGATATC 826
RESULT 33
ABLO7182/c
ID      ABL07182 standard; cDNA; 4406 BP.
XX
AC      ABL07182;
XX
DT      26-MAR-2002 (first entry)
XX
DE      Drosophila melanogaster expressed polynucleotide SEQ ID NO 16028.
XX
KW      Drosophila; developmental biology; cell signalling; insecticide;
XX
KW      pharmaceutical; gene; ss.
XX
OS      Drosophila melanogaster.

```


XX SQ Sequence 422 BP; 95 A; 125 C; 121 G; 81 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.82e-27	Length:	422
Score:	324.00	Matches:	62
Percent Similarity:	78.95%	Conservative:	28
Best Local Similarity:	54.39%	Mismatches:	24
Query Match:	33.68%	Indels:	0
DB:	9	Gaps:	0

US-10-627-571-2 (1-188) x ACH48309 (1-422)

QY 6 PheasnSerLysaenLeuAlaValGlnAlaGlnLysLysLeuGlyLysMetValSer 25
 DB 77 TTCAAGCTCAAGAGCCCTGGCACTGCAAGCAAGAGAGAGCTGAGTAAAGTGGCGGT 136
 QY 26 LysSerLeuAlaThrThrLeuLysAspThrSerSerGluValLeuAspGluLeuTyr 45
 DB 137 CGCTCTGTGGCTACTCTCTCATGATGACACAGAGCTAGAGTGTGCTTAAATGAGCTTAC 196
 QY 46 ArgValThrArgGluTyrThrGlnAsnLysLysGlnAlaGluLysLysLysAsnLeu 65
 DB 197 CGTGTGCCAAGAGGTACAGCACAGCCGGGCCAGGCCAGCGCGTATCAAGACCTG 256
 QY 66 IleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGln 85
 DB 257 ATCAAAATGGCCATCAAGGTGGTGTCTGCTGCAATGAGCTCTTGGCCCCAGTGAG 316
 QY 86 LeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAlaMetThrValValSer 105
 DB 317 CTGGCCCTGGCTACCCCTTTCGCCAAGAGCTGGCGGAGGTGCTCATGACGCACTTAC 376
 QY 106 PheHisGlnValAspTyrThrPheAspArgAsnValLeuSer 119
 DB 377 TTGTGTGAGGTAGACTTCACTTCAAGCGTGTGCTTGTGCT 418

RESULT 35

ID ACH76016 standard; DNA; 500 BP.

XX ACH76016;

XX AC 29-JUN-2004 (first entry)

XX DT 29-JUN-2004 (first entry)

XX DE Human genome derived single exon probe #9211.

XX KM Human; probe: ss; gene expression; single exon probe; microarray;

XX KW alternative splicing event; genomic alteration.

XX OS Homo sapiens.

XX PN US2003194704-A1.

XX PD 16-OCT-2003.

XX PF 03-APR-2002; 2002US-00029386.

XX PR 03-APR-2002; 2002US-00029386.

XX PA (PENN/) PENN S G.

XX PA (RANK/) RANK D R.

XX PA (HANZ/) HANZEL D K.

XX PI Penn SG, Rank DR, Hanzel DK,

XX DR WPI; 2004-119264/12.

XX PT New human genome-derived single exon nucleic acid probes useful for human

XX PT gene expression analysis, for identifying or characterizing alternative

XX PT splicing events, for assessing genomic alterations or as tools for

XX PT surveying tissues.

PS Claim 15; SEQ ID NO 9211; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene

CC expression, comprising any of the 27,400 fully defined nucleotide

CC sequences in the specification, or their complements or fragments, and

CC encoding at least 8 amino acids of any of the 6888 amino acid sequences

CC fully defined in the specification. The probe is a single exon probe that

CC hybridizes under high stringency conditions to a nucleic acid molecule

CC expressed in human cells or tissues. Also included are a spatially-

CC addressable set of single exon nucleic acid probes for measuring human

CC gene expression (comprising a plurality of single exon nucleic acid

CC probes cited above, where each of the plurality of probes is separately

CC and addressably isolatable or amplifiable from the plurality), a single

CC exon microarray for measuring human gene expression, a method of

CC measuring human gene expression, a vector comprising the single exon

CC probe cited above, an ORF-encoded peptide comprising at least 8

CC contiguous amino acids of any of the above-mentioned amino acid

CC sequences (optionally with conservative amino acid substitutions), an

CC isolated antibody that binds specifically to a peptide cited above,

CC methods of selling and/or licensing single exon probes or microarrays to

CC a customer desiring to measure gene expression, a method of providing

CC human gene expression data by subsequence, and a computer-readable

CC storage medium which contains a database having a plurality of records

CC (each record including data on the expression of a single exon probe

CC cited above. The probe, methods and apparatus are useful in gene

CC expression analysis. The probes may be used as tools for surveying

CC tissues to detect the presence of expressed messages that contain their

CC specific exon, or in constructing genome-derived single exon microarrays.

CC In addition, the probes are used in identifying and characterizing

CC alternative splicing events, in detecting and characterizing gross

CC alterations in the genomic locus that includes their exon, in assessing

CC smaller genomic alterations, in priming the synthesis of nucleic acids,

CC or in expressing the ORF-encoded peptide. The present sequence is a human

CC single exon probe of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?DocID=20030194704

XX SEQ Sequence 500 BP; 164 A; 71 C; 107 G; 158 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.25e-25	Length:	500
Score:	304.00	Matches:	55
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	31.60%	Indels:	0
DB:	12	Gaps:	0

US-10-627-571-2 (1-188) x ACH76016 (1-500)

QY 134 GlnArgHisLeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSer 153
 DB 3 CAGCGCCACTCTACTCCCAAGTACATGACCGGTTAATATGTGTTGATCATTTTCA 62
 QY 154 AspCysGluPheLeuAlaIleuTyrAsnProPheGlyAsnPhelysProHisLeuGln 173
 DB 63 GATTGTGAATTTTGGCTGCTCTGTATATCTTTTGGGAATTTTAAACCCCACTTAA 122
 QY 174 LysLeuCysAspGlyIleAsnLysMetLeuAspGluGlnAsnIle 188
 DB 123 AAACATGTGATGTATCAACAATAATGTTGAGAGAGAACATA 167

RESULT 36

ID AAH04721 standard; cDNA; 714 BP.

XX AAH04721;

XX AC 26-JUN-2001 (first entry)

XX DT 26-JUN-2001 (first entry)

XX DE Human cDNA clone (5'-primer) SEQ ID NO:1556.

XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

```

XX OS Homo sapiens.
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX PI WPI; 2001-318749/34.
XX DR
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 1, SEQ ID NO 1556; 2537bp + Sequence Listing; English.
XX
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification; where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX
XX SQ Sequence 714 BP; 113 A; 246 C; 203 G; 148 T; 0 U; 4 Other;
XX
XX Alignment Scores:
XX Pred. No.: 1,566-23 Length: 714
XX Score: 292.00 Matches: 50
XX Percent Similarity: 81.188 Conservative: 19
XX Best Local Similarity: 58.828 Mismatches: 16
XX Query Match: 30.354 Indels: 0
XX DB: 4 Gaps: 0
XX
XX US-10-627-571-2 (1-188) x AAH04721 (1-714)
XX
XX QY 104 ValSerPheHisGlnValAspTyrThrPheAspArgAsnValIleuSerArgIleuLeuAan 123
XX DB 6 GTCAGCTTCACACGAGTGAACCTTCGACCGCGCGTGGCCGCCCGGAGCTGCCTC 65
XX
XX QY 124 GlnCysArgGluMetLeuHisGlnIleIleGlnArgHisIleuThrAlaIysSerHisGly 143
XX DB 66 GAGTCCCGGACCTGCTGACCCAGGCGCGTGGCTCCCGACCTTGACCGCAAGTCCACGGCC 125
XX
XX QY 144 ArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIleuTyrAan 163

```

```

DB 126 CCGATCAACACAGCTTTGGCCACCTACCGACCTGCTGCGCTACAGCC 185
QY 164 PropheGlyAsnPhelyeProHisIleuGlnIlyLeuCyAspGlyTleAsnIyMetIeu 183
DB 186 CCGCGCGAGCCCTACCGCTCCACCTGCGAGATCTCGAGGCGCTGCGCGATGCTG 245
QY 164 AspGluGluAenIle 188
DB 246 GACGAGGCGAGCCTC 260
RESULT 37
AAH18715
ID AAH18715 standard; cDNA; 1602 BP.
AC AAH18715;
XX
XX DT 26-JUN-2001 (first entry)
XX
XX DE Human cDNA sequence SEQ ID NO:18986.
XX
XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
XX OS Homo sapiens.
XX
XX PN EP1074617-A2.
XX PD 07-FEB-2001.
XX PF 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX
XX PA (HELI-) HELIX RES INST.
XX
XX PI Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX PI WPI; 2001-318749/34.
XX
XX DR
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 8; SEQ ID NO 18986; 2537bp + Sequence Listing; English.
XX
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification; where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention

```

```
XX      SO      Sequence 1602 BP; 334 A; 458 C; 491 G; 319 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.:      4,52e-23      Length:      1602
Score:          292.00      Matches:      50
Percent Similarity: 81.18%      Conservative: 19
Best Local Similarity: 58.82%      Mismatches:  16
Query Match:     30.35%      Indels:      0
DB:              4      Gaps:      0

US-10-627-571-2 (1-188) x AAH18715 (1-1602)

QY      104  ValSerPheHsiGlnValAspGlyTThPheAspArgAsnValLeuSerArgLeuLeuAsn 123
        6   GTGAGCTTCCACGAGGAGGAGCTTCACTTCGACCGGCGGCGGCGGCGGCTGCTC 65
        124  GTCyAArgGluMetLeuHsiGlnIleIleGlnAArgHsiLeuThraIalysSerHisGly 143
        66  GAGTGGCGGACCTGCTGACACGAGCGCGTGGGTCCTCCACCTGACCGCAAGTCCACAGGC 125
QY      144  ArgValAsnAsnValPheAspHisPheSerAspCySGluPheLeuAlaIalLeuTyzAsn 163
        126  CGGATCAACACGAGTGTGGGACACCTAGCCGACCTGACCTTCGCTGGGCTTACGGC 165
QY      164  ProPheGlyAsnPhelyPProHsiLeuGlnIlyLeuCyAspGlyIleAsnIlyMetLeu 183
        186  CCGCGCCGAGCCCTACGCTCCGACCTGCGAGATCTGCGAGGCGCTGCGGCGGATGCTG 245
        184  AspGluGluAsnIle 188
        246  GACGAGGGCAGCCTC 260
DB
QY      184  AspGluGluAsnIle 188
        246  GACGAGGGCAGCCTC 260
DB

RESULT 38
ID      AAK56622
ID      AAK56622 standard; cDNA; 441 BP.
XX
AC      AAK56622;
XX
DT      06-NOV-2001 (first entry)
XX
DE      Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1682.
XX
KW      Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
        cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS      Homo sapiens.
XX
PN      W0200157182-A2.
XX
PD      09-AUG-2001.
XX
PF      17-JAN-2001; 2001WC-US001354.
XX
PR      31-JAN-2000; 2000US-0179065P.
PR      04-FEB-2000; 2000US-0180628P.
PR      24-FEB-2000; 2000US-0184664P.
PR      02-MAR-2000; 2000US-0186350P.
PR      16-MAR-2000; 2000US-0189874P.
PR      17-MAR-2000; 2000US-0190076P.
PR      18-APR-2000; 2000US-0198133P.
PR      19-MAY-2000; 2000US-0205515P.
PR      07-JUN-2000; 2000US-0209467P.
PR      28-JUN-2000; 2000US-0214886P.
PR      30-JUN-2000; 2000US-0215135P.
PR      07-JUL-2000; 2000US-0216647P.
PR      07-JUL-2000; 2000US-0216880P.
PR      11-JUL-2000; 2000US-0217487P.
PR      14-JUL-2000; 2000US-0218290P.
PR      26-JUL-2000; 2000US-0220963P.
PR      26-JUL-2000; 2000US-0220964P.
PR      14-AUG-2000; 2000US-0224518P.
```

```
PR      14-AUG-2000; 2000US-0224519P.
PR      14-AUG-2000; 2000US-0225213P.
PR      14-AUG-2000; 2000US-0225214P.
PR      14-AUG-2000; 2000US-0225266P.
PR      14-AUG-2000; 2000US-0225267P.
PR      14-AUG-2000; 2000US-0225268P.
PR      14-AUG-2000; 2000US-0225270P.
PR      14-AUG-2000; 2000US-0225477P.
PR      14-AUG-2000; 2000US-0225757P.
PR      14-AUG-2000; 2000US-0225758P.
PR      14-AUG-2000; 2000US-0225759P.
PR      18-AUG-2000; 2000US-0226279P.
PR      22-AUG-2000; 2000US-0226681P.
PR      22-AUG-2000; 2000US-0226688P.
PR      22-AUG-2000; 2000US-0227182P.
PR      23-AUG-2000; 2000US-0227009P.
PR      30-AUG-2000; 2000US-0228924P.
PR      01-SEP-2000; 2000US-0229287P.
PR      01-SEP-2000; 2000US-0229343P.
PR      01-SEP-2000; 2000US-0229344P.
PR      01-SEP-2000; 2000US-0229345P.
PR      05-SEP-2000; 2000US-0229509P.
PR      06-SEP-2000; 2000US-0229513P.
PR      06-SEP-2000; 2000US-0230437P.
PR      06-SEP-2000; 2000US-0230438P.
PR      08-SEP-2000; 2000US-0231422P.
PR      08-SEP-2000; 2000US-0231433P.
PR      08-SEP-2000; 2000US-0231444P.
PR      08-SEP-2000; 2000US-0231413P.
PR      08-SEP-2000; 2000US-0231414P.
PR      08-SEP-2000; 2000US-0232080P.
PR      08-SEP-2000; 2000US-0232081P.
PR      12-SEP-2000; 2000US-0231968P.
PR      14-SEP-2000; 2000US-0232397P.
PR      14-SEP-2000; 2000US-0232398P.
PR      14-SEP-2000; 2000US-0232399P.
PR      14-SEP-2000; 2000US-0232400P.
PR      14-SEP-2000; 2000US-0232401P.
PR      14-SEP-2000; 2000US-0233063P.
PR      14-SEP-2000; 2000US-0233064P.
PR      14-SEP-2000; 2000US-0233065P.
PR      21-SEP-2000; 2000US-0234423P.
PR      21-SEP-2000; 2000US-0234474P.
PR      25-SEP-2000; 2000US-0234997P.
PR      25-SEP-2000; 2000US-0234998P.
PR      25-SEP-2000; 2000US-0235484P.
PR      27-SEP-2000; 2000US-0235834P.
PR      27-SEP-2000; 2000US-0235836P.
PR      27-SEP-2000; 2000US-0236327P.
PR      29-SEP-2000; 2000US-0236367P.
PR      29-SEP-2000; 2000US-0236368P.
PR      29-SEP-2000; 2000US-0236369P.
PR      29-SEP-2000; 2000US-0236370P.
PR      02-OCT-2000; 2000US-0236802P.
PR      02-OCT-2000; 2000US-0237037P.
PR      02-OCT-2000; 2000US-0237038P.
PR      02-OCT-2000; 2000US-0237039P.
PR      02-OCT-2000; 2000US-0237039P.
PR      02-OCT-2000; 2000US-0237040P.
PR      13-OCT-2000; 2000US-0239835P.
PR      13-OCT-2000; 2000US-0239837P.
PR      20-OCT-2000; 2000US-0240960P.
PR      20-OCT-2000; 2000US-0241221P.
PR      20-OCT-2000; 2000US-0241785P.
PR      20-OCT-2000; 2000US-0241786P.
PR      20-OCT-2000; 2000US-0241787P.
PR      20-OCT-2000; 2000US-0241808P.
PR      20-OCT-2000; 2000US-0241809P.
PR      20-OCT-2000; 2000US-0241826P.
PR      01-NOV-2000; 2000US-0244617P.
PR      08-NOV-2000; 2000US-0246474P.
PR      08-NOV-2000; 2000US-0246475P.
PR      08-NOV-2000; 2000US-0246476P.
PR      08-NOV-2000; 2000US-0246477P.
```

PR	08-NOV-2000;	2000US-0245478P.
PR	08-NOV-2000;	2000US-0245523P.
PR	08-NOV-2000;	2000US-0245524P.
PR	08-NOV-2000;	2000US-0245525P.
PR	08-NOV-2000;	2000US-0245526P.
PR	08-NOV-2000;	2000US-0245527P.
PR	08-NOV-2000;	2000US-0245528P.
PR	08-NOV-2000;	2000US-0245532P.
PR	08-NOV-2000;	2000US-024609P.
PR	08-NOV-2000;	2000US-024610P.
PR	08-NOV-2000;	2000US-024611P.
PR	08-NOV-2000;	2000US-0246113P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251088P.
PR	05-DEC-2000;	2000US-02556719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	08-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254097P.
PR	05-JAN-2001;	2001US-0255678P.
PA	(HMA-) HUMAN GENOME SCI INC.	
XX		
XX		
PI	Rosen CA,	Barash SC, Ruben SM;
XX		
DR	WPI; 2001-483426/52.	
DR	P-PSDB; AAM83841.	
XX		
PT	Nucleic acids encoding human immune/haematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
XX		
PS	Claim 1; SEQ ID NO 1682; 3071pp + Sequence Listing; English.	
PS		
CC	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
CC	amino acid sequences given in AAM82170 to AAM91321. (I) have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patient's own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoietic-related diseases, especially	
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
CC	to AAK87694 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	

CC	represent sequences used in the exemplification of the present invention					
XX	Sequence	441 BP; 99 A; 119 C; 127 G; 89 T; 0 U; 7 Other;				
SQ						
	Alignment Scores:					
	Pred. No.:	1 47e-22	Length:	441		
	Score:	281.00	Matches:	60		
	Percent Similarity:	74.36%	Conservative:	27		
	Best Local Similarity:	51.28%	Mismatches:	24		
	Query Match:	29.21%	Indels:	7		
	DB:	4	Gaps:	1		
	US-10-627-571-2 (1-188) x AAK56622 (1-441)					
OY	6 PheamserLysasndeuAlavalGlnalaglnLysileLeugLyLysmetValSer	25				
Db	89 TTACAGCTCAAGAAGCCTGGGACTGCACAGCAGAAGAAGCTACTAGTAAGATGGCGGT	14				
OY	26 LysSerIleAlaThrThrIleuIleaspAspThrSerSerGIuValIleuaspGIuLeuTyr	45				
Db	149 CGCCTGTGGCTCATCTCTTCATAGATGAGACAAGACAGTAGTGATCATGATGACTTAC	20				
OY	46 ArgValThrArgIuLurTrThrgInaenLysylsGIuaIagluLysIleLysasndeu	65				
Db	209 CGTGTGTCCAAAGAGTACAGCACACCGCCGCCAGCCGACGGGTGATCAGGACCTG	26				
OY	66 IleIleThrValIleLysIleuAlaIleLeuTyrArAsnAsnGlnPheasnGlnaspGu	85				
Db	269 ATCAAAGTGGCCATCAAGGTGGCTGTCTGCACACCGAATGGCTCTTGGCCCCAGTGAG	32				
OY	86 LeuAlaLeuMetCIuLysPheLysIlyslsValHIsGlnLeuAlaMetThrValValSer	10				
Db	329 CTGGCCTGTGGCTAACCGGCTTTGGCCAGARCTGGCGCAAGGTGCCATGAMGGCATTGAGC	38				
OY	106 PheHisGlnValaspTyrThrPheaspAspValLeuSerArgLeuLeu	122				
Db	389 TTGGT-GAAGTNGATTTCAC-----TCGAAGCTGCTG	420				
	RESULT 39					
ID	AAc10352 standard, cDNA; 340 BP.					
XX	AAc10352;					
AC						
XX	06-OCT-2000 (first entry)					
DT						
XX						
DE	Human secreted protein 5' EST, SEQ ID NO: 14427.					
XX						
KW	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;					
XV	gene therapy; chromosome mapping; ss.					
OS	Homo sapiens.					
XX						
PN	EPI033401-A2.					
XX						
PD	06-SEP-2000.					
PE	21-FEB-2000; 2000EP-00200610.					
XX						
PR	26-FEB-1999; 99US-0122487P.					
PA	(GEST) GENSET.					
XX						
PI	Dumas Milne Edwards J, Duclert A, Giordano J;					
XX						
DR	WPI; 2000-500381/45.					
PT	New nucleic acid that is a 5' expressed sequence tag (5' EST) for					
PT	obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for					
PT	diagnostic, forensic, gene therapy and chromosome mapping procedures.					
XX						
PS	Claim 1; SEQ ID NO 14427; 71bp + Sequence Listing; English.					
XX						

CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNA or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors

XX SQ Sequence 340 BP; 92 A; 79 C; 85 G; 79 T; 0 U; 5 Other;

Alignment Scores:

Pred. No.:	3,44e-19	Length:	340
Score:	250.00	Matches:	52
Percent Similarity:	96.36%	Conservative:	1
Best Local Similarity:	94.55%	Mismatches:	2
Query Match:	25.99%	Indels:	0
DB:	3	Gaps:	0

US-10-627-571-2 (1-188) x AAC10352 (1-340)

QY 2 AlAThAspYAlPheAsnSerLYsAsnLeuAlaValGlnAgluLYsLYleLeuGly 21

Db 174 GCACAGATGCTCTTAATTCAAAACCTGGCGCKTCAGGACCAAAABAAGATCTTGCT 233

QY 22 LyMeTValSerLYsSerLYleAlaThrThLeuILeAspAspThrSerSerGluValLeu 41

Db 234 AAATGTGTGTCAAATCATCGCCACCACTTAATAGACGACCAAGTGTGAGTGCTG 293

QY 42 AspGluLeuTYrArgValThrArgGluTYrThGlnAsnLYsLYs 56

Db 294 GATGAGCTCTACAGAGTACGAGGAGTACCCCAAAACAGAG 338

RESULT 40

AA192435/c

AA192435 standard; cDNA; 400 BP.

AC AA192435;

DT 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 12495.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

XX tissue growth factor; immunomodulatory; cancer; leukaemia;

XX nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US004927.

PR 28-FEB-2000; 2000US-00515126.

PA 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX P-PSDB; AAO12504.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing

PT and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 1; SEQ ID NO 12495; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AA179941-AA193841) and

XX the encoded proteins (AA000010-AA013910) that exhibit activity relating to

XX cytokine, cell proliferation or cell differentiation or which may induce

XX production of other cytokines in other cell populations. The

XX polynucleotides and polypeptides are useful in gene therapy, vaccines or

XX peptide therapy. The polypeptides have various cytokine-like activities,

XX e.g. stem cell growth factor activity, haematopoiesis regulating

XX activity, tissue growth factor activity, immunomodulatory activity and

XX activation/inhibition activity and may be useful in the diagnosis and/or

XX treatment of cancer, leukaemia, nervous system disorders, arthritis and

XX inflammation. Note: The sequence data for this patent did not form part

XX of the printed specification, but was obtained in electronic format

XX directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 400 BP; 80 A; 122 C; 114 G; 84 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	4.2e-14	Length:	400
Score:	206.00	Matches:	38
Percent Similarity:	85.25%	Conservative:	14
Best Local Similarity:	62.30%	Mismatches:	9
Query Match:	21.41%	Indels:	0
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x AA192435 (1-400)

QY 2 AlAThAspYAlPheAsnSerLYsAsnLeuAlaValGlnAgluLYsLYleLeuGly 21

Db 184 GCCATGGACACCTTTCAGACCAAGAGCTGGCTCTCAGGCGGAGAAAGCTCTGAGT 125

QY 22 LyMeTValSerLYsSerLYleAlaThrThLeuILeAspAspThrSerSerGluValLeu 41

Db 124 AACATGGCTCCAAAGGACGTGGCGCTGTGTGATGACCCAGCATGTGAGTGCTG 65

QY 42 AspGluLeuTYrArgValThrArgGluTYrThGlnAsnLYsLYs 61

Db 64 GATGAGCTGTACCGGCCACCGAGGAGTTCACCGCGAGCGGACGAGAGGCCAGAGAG 5

QY 62 Ile 62

Db 4 CTC 2

RESULT 41

AAC10359

AA10359 standard; cDNA; 252 BP.

AC AA10359;

DT 06-OCT-2000 (first entry)

XX Human secreted protein 5' EST, SEQ ID NO: 14434.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

XX gene therapy; chromosome mapping; ss.

OS Homo sapiens.

PN BP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-00200610.

PR 26-FEB-1999; 99US-0122487P.

PA (GEST) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 14434; 71bp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A⁺ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 252 BP; 57 A; 88 C; 60 G; 46 T; 0 U; 1 Other;
XX
Alignment Scores:
Pred. No.: 7 92e-10 Length: 252
Score: 166.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.26% Indels: 0
DB: 3 Gaps: 0
US-10-627-571-2 (1-188) x AAC10359 (1-252)
QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLyIleLeu 20
DB 148 ATGGCCACGATGCTCTTAATTCGAAACCTGCGCGTTCAGGCACAAAGAGCTTG 207
QY 21 GlyLyMetValSerLySerIleAlaThrThleuIleAspAsp 35
DB 208 GGTAAATGGTGTCCAAATCCATGCGCACCACTTAATGACGAC 252
RESULT 42
ACH78513
ID ACH78513 standard; DNA; 543 BP.
XX
XX ACH78513;
XX
XX 29-JUL-2004 (first entry)
XX
XX Human genome derived single exon probe #11708.
XX
XX Human; probe; ss; gene expression; single exon probe; microarray;
XX alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
XX US2003194704-A1.
XX
XX 16-OCT-2003.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX 03-APR-2002; 2002US-00029386.
XX
XX (PENN/) PENN S G.
XX (RANK/) RANK D R.
XX (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK;
XX
XX WPI; 2004-119264/12.
XX

XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 15; SEQ ID NO 11708; 80bp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 688 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human
CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the database having a plurality of records
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 543 BP; 137 A; 128 C; 119 G; 159 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2.4e-07 Length: 543
Score: 148.00 Matches: 24
Percent Similarity: 88.64% Conservative: 15
Best Local Similarity: 54.55% Mismatches: 5
Query Match: 15.38% Indels: 0
DB: 12 Gaps: 0
US-10-627-571-2 (1-188) x ACH78513 (1-543)
QY 143 GlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeuTyr 162
DB 3 GGCGGATACACCGATCTTTAAACCACTTGGCGATGGAGTCTCTCCACCCCTCAT 62
QY 163 AsnProPheGlyAsnPhelySProHisLeuGlnLyLeuCyAspGlyIleAsnLySme 182
DB 63 AGTCTGATGAGACACTGTAGCGCCCAACTCAAGAGATTGTGCAAGATCATTAAGTTG 122
QY 183 LeuAspGluGlu 186
DB 123 CTAGATGAGAAA 134
RESULT 43
ADP04796
ID ADP04796 standard; cDNA; 1329 BP.
XX
XX ADP04796;
XX

XX	29-UUL-2004 (first entry)	
DT		
XX	Sea squirt cDNA with tissue specific expression in development Seg 391.	
XX		
XX	gene; ss; sea squirt; regeneration medicine; gene therapy;	
KV	cell proliferation; differentiation; reproduction;	
KW	environmental measurement; water survey.	
XX		
OS	Ciona intestinalis.	
XX		
PN	JP2004057129-A.	
PD		
XX	26-FEB-2004.	
XX		
XX	31-UUL-2002; 2002JP-00222593.	
PF		
XX		
PR	31-UUL-2002; 2002JP-00222593.	
XX		
PA	(KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.	
XX		
DR	WPI, 2004-287079/27.	
DR	P-PSDB; ADP04797.	
XX		
PT	Novel gene cluster which is specifically expressed in tissue or organ	
PT	during developmental phase of sea squirt, useful for elucidation of	
PT	mechanism of development of tissue or organ of sea squirt.	
XX		
PS	Claim 2; SEQ ID NO 391; 1846bp; Japanese.	
XX		
CC	This invention relates to novel genes and the encoded proteins thereof	
CC	that are derived from the sea squirt Ciona intestinalis. Specifically, it	
CC	refers to those genes that are expressed in the tissues or organs of the	
CC	sea squirt during its developmental phase. The present invention	
CC	describes the identification of these genes as useful for elucidation of	
CC	the mechanism of development and hence for developing regeneration	
CC	medicines and gene therapy techniques. Accordingly, they can be used in	
CC	the research of various genetic diseases, as well as the analysis of cell	
CC	proliferation, differentiation and reproduction. Furthermore, such	
CC	compositions can be useful for environmental measurements and water	
CC	surveys, particularly for sea water surveys, and also for the preparation	
CC	of transformed sea squirt for improving edibility of sea squirt such as	
CC	Halocynthia roretzi. This polynucleotide sequence is a sea squirt cDNA	
CC	sequence that exhibits tissue specific expression during development,	
CC	given in an exemplification of the invention.	
XX		
SO	Sequence 1329 BP; 596 A; 183 C; 280 G; 270 T; 0 U; 0 Other;	
	Alignment Scores:	
	Pred. No.:	0.0456
	Score:	106.00
	Percent Similarity:	44.51%
	Best Local Similarity:	24.73%
	Query Match:	11.02%
	DB:	12
		Gaps: 8
US-10-627-571-2 (1-188) x ADP04796 (1-1329)		
OY	7 AaenSerlySaenLeuAlaValGlnAlaGlnlySylleuGlnlySmetValSerlyS	26
DB	242 AATGAAGAAAACCTGCTGCACAAAGCCGAGAAACA-----	277
OY	27 SerileAlaThrThleuLleAspAspTrSerSerGluValleuAspGluLeuYrYrG	46
DB	278 -----AAGAAATTAATAAGATGAAGAAAGAAAGAAAGTTCGAAGACGAG-----	319
OY	47 ValThrArgGluTrYThrGlnAsnlySylGlnAlaGlnlySylleuSylleuSAsnleuile	66
DB	320 GTT---CGTAAACAAATCGATACCAAGAAAGAAAGAAAGAAAGAAAGAAAGTCTTAA	376
OY	67 LysThrValleuLleuLleuLleuYrYrArgAsnAsnGlnPhe-----	Asn 82
DB	377 AAGGAAGACAAATTAAGAGCTATTGTGAAGATCATTAATTCCTTCGTGAATCTGAA	436

QY	83	GINASPGILuleuAlaleuMeGLuysPheIySylvAlHisIleuAlMeTthr	102
DB	437	GAGCAAGACTTAAATGATGGAA-----CAATGATCATCTTGTTGGAA	484
QY	103	Val-----ValSerPheHisIleValAspIyrThrPheAsp-----	114
DB	485	GTCAATCTTATTTCTTCGACAGAAATTAAACAATGATGATAGGATGACAGATGTACTGAA	544
QY	115	-----ArgAsnValIleuSerArgLeuIleuAsnGluCysArgGluMetLeuHisGlnIle	132
DB	545	AGTGCACGAATATATATTGAACAAAGCTTCCAAACTTCGGCAACAAGTTGAAACAAG	604
QY	133	IIeGIaNGHlslenThrAlalysSerHisSGIyArgValAsnAsnValPheAsp-----	150
DB	605	AAACGAAACATTGGGAAGACGCTTCCAAAGTCAAGAGATGACGCTGTGAATGGAAGTTCA	664
QY	151	-----HisPheSerAspCysGluPheIleuAlaAlaIleuTyrAsn	163
DB	665	AGTGCAGAGAGAAACATCGAGTTACAGTATATTAAATCTTGATTAAAGCAGTCAAT	724
QY	164	ProPhe	165
DB	725	CTCTTT	730
RESULT 44			
ID	AAA70229	standard; DNA, 2943 BP.	
AC	AAA70229;		
XX	07-NOV-2000	(first entry)	
DE	Plasmodium falciparum chromosome 2 related DNA sequence SEQ ID NO:362.		
XX	Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine;		
KW	antimalarial; malaria; protozoasidae; infection; insecticide; ds.		
XX	Plasmodium falciparum.		
OS	WO200025728-A2.		
XX	11-MAY-2000.		
PD	05-NOV-1999;	99WO-US026796.	
XX	05-NOV-1998;	98US-0107131P.	
PR	(HOFF/) HOFFMAN S.		
XX	(CARU/) CARUCCI D.		
PA	(GARD/) GARDNER M.		
PA	(VENT/) VENTER J C.		
XX	Hoffman S, Carucci D, Gardner M, Venter JC;		
PI	WPI; 2000-365347/31.		
XX	Proteins encoded by chromosome 2 of the human malarial parasite,		
PT	Plasmodium falciparum, useful as antimalarial vaccines and in the		
PT	diagnosis of P.falciparum infection.		
XX	Disclosure; Page 549-550; 577pp; English.		
XX	The present invention describes proteins and their fragments (I) encoded		
CC	by chromosome 2 of the human malarial parasite, Plasmodium falciparum.		
CC	Also described are: (I) nucleotide sequences (II) encoding (I); and (2)		
CC	vaccines against P. falciparum infection comprising (I) or (II). (I) and		
CC	(II) are useful for the development of vaccines against P. falciparum		
CC	infection. (I) and polyclonal antisera or a monoclonal antibody raised to		
CC	immunogens comprising the sequences of (I), are useful in the detection		
CC	of infection with P. falciparum. Furthermore, (I) (especially when they		
CC	are rifins or secreted or membrane proteins) can aid the identification		
CC	of drugs to treat or prevent P. falciparum infection, or they can be used		

CC to identify drug resistance in *P. falciparum*. Sequencing of the
 CC plasmodium chromosome 2 and the subsequent identification of proteins
 CC encoded by it will help to expand our understanding of parasite biology,
 CC a process hampered by the complexity of the parasitic lifecycle, and
 CC provide new targets for vaccine and drug development. Parasite resistance
 CC to drugs and mosquito resistance to insecticides have led to a resurgence
 CC of malaria in many parts of the world, and there is a pressing need for
 CC vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352
 CC represent nucleotide and protein sequences given in the present
 CC invention, but which are not specifically mentioned within the
 CC specification

XX Sequence 2943 BP; 1587 A; 166 C; 520 G; 670 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0 219	Length:	2943
Score:	104.00	Matches:	42
Percent Similarity:	43.81%	Conservative:	43
Best Local Similarity:	21.65%	Mismatches:	79
Query Match:	10.81%	Indels:	30
DB:	3	Gaps:	9

US-10-627-571-2 (1-188) x AAA70229 (1-2943)

```

QY 1 MetAlaThrAspValPheAsnSerIySaenLeuAlaValGlnAlaGlnIySylleu 20
    |||  ::  ::  |||  |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 1588 ATGATATATCGAACAAATATACAAAAAG-----ATACAGAAAGAAAAGATATGCTA 1638
QY 21 GlyIySmetValSerIySerIleAlaThrIleuIleAspAspThrSerSerGluVal 40
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 1639 AATATATATATATATAAGT-----ATGATCTGACACATATAAATCA 1680
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAspIySylleuGluVal 60
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 1681 ATTGAAAGAAATGCAAGAA-----GAAATATAAAGAACTT---AAG 1719
QY 61 IySylleuAsnLeuIleuIySylleuValIleuIySylleuValIleuTyArgSnaenGln 80
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 1720 AAGCTGAAAATGATATGATATGATCTACAGTCACAAATTTTATTAAGGAAATGAA 1779
QY 81 PheAsnGluAspGluLeuAlaLeuMetGluIySylleuPheIySylleuValHis 97
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 1780 AAGCAGATGCAAGAA-----AAAGTGAGAGAAATATATAAATTATTAAGCAAGAAAGAT 1833
QY 98 GlnLeuAlaMetThrValIySerPheHisGlnValAspTyThrPheAspArgAsnVal 117
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 1834 CAAAGAACTTAAATATATTTACAGAAATATGATGAAAGAAATGAAATTAAGAA 1893
QY 118 LeuSerArgLeuLeuAsnGluCySArgGluMetLeuHisGln-----IleIleGlnArg 135
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 1894 ATCGAAGATATATCGTAAATGATTTGTGAAGAAAGTTGACAGCACTAAATATATATA 1953
QY 136 HisLeuThrAlaIySerHisGlyArgValAsnAsn-----ValPheAspHisPheSer 153
    |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 1954 AAATTAACCAACCGCAAGATATGCGCAATATATATATATATGCTCATGAT----- 2004
QY 154 AspCyGluPheLeuAlaIleuTyArgProPheGluIyAsnPheIySylleuGln 173
    |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 2005 -----GAAATTTTAAAGAAAGAAAGTAAATAAATTAACGATTTGTAAGATATGGA 2058
QY 174 IySleuCySAspGluIleAsnIySmetLeuAspGluGluAsn 187
    |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 2059 AAGAAAGAAAGAGAAATTAATTAAGCTCGTGAAGAAAGAAAGAACT 2100

```

RESULT 45

AAA95803

ID AAA95803 standard; DNA; 588 BP.

AC AAA95803;

DT 09-MAR-2001 (first entry)

XX Tomato E-196 nucleotide sequence.

XX Tomato; E-196; MAR binding filament-like protein 1; MFPI;
 KW matrix attachment region; MAR; NEMFPI-2; anchor protein; ds.
 XX Lycopersicon esculentum.

Key	Location/Qualifiers
FT CDS	2..588
FT	/tag= a
FT	/partial
FT	/product= "E-196"

PN WO20061615-A2.

PD 19-OCT-2000.

PF 12-APR-2000; 2000WO-US009723.

PR 12-APR-1999; 99US-0128900P.

PA (DUPO) DU FONT DE NEMOURS & CO E I.

PI Harder PA, Weier I;

DR WPI; 2000-679464/66.

DX P-PSDB; AAB21229.

PT Nucleic acid fragments from tobacco, corn, soybean and rice, encoding
 PT proteins that are homologs to the MAR binding filament-like protein 1
 PT (MFPI) useful for development of novel phenotypes.

PS Example 2; Page 49; 62pp; English.

CC The present sequence encodes E-196, a polypeptide fragment isolated from
 CC the tomato. E-196 is a subfragment of the matrix attachment region (MAR)
 CC binding filament-like protein 1 (MFPI). A recombinant fusion protein
 CC comprising an N-terminal 6-histidine tag fused to E-196 was used to raise
 CC rabbit polyclonal antibody a288 against E-196. This was used to detect
 CC proteins with antigenic similarity to MFPI in other plant species. MFPI
 CC has features of a novel anchor protein that most likely connects
 CC chromatin via MAR DNA with the nuclear envelope and nuclear filament
 CC proteins. MFPI nuclear acids and proteins may be used to better
 CC understand the mechanisms underlying this process so that the attachment
 CC of transgenes to the nuclear matrix may be used routinely to improve gene
 CC expression. They may be used to study MFPI expression, leading to the
 CC creation of novel developmental phenotypes that may be beneficial for
 CC crop growth and development. In addition, if the reduction in expression
 CC of one of the genes leads to a growth or developmental defect in the
 CC plant, this gene can be used as a novel herbicide target

SQ Sequence 588 BP; 223 A; 86 C; 142 G; 137 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0 187	Length:	588
Score:	96.50 <td>Matches:</td> <td>43 </td>	Matches:	43
Percent Similarity:	40.21% <td>Conservative:</td> <td>33 </td>	Conservative:	33
Best Local Similarity:	22.75% <td>Mismatches:</td> <td>90 </td>	Mismatches:	90
Query Match:	10.03% <td>Indels:</td> <td>23 </td>	Indels:	23
DB:	3 <td>Gaps:</td> <td>6 </td>	Gaps:	6

US-10-627-571-2 (1-188) x AAA95803 (1-588)

```

QY 1 MetAlaThrAspValPheAsnSerIySaenLeuAlaValGlnAlaGlnIySylleu 20
    ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 53 TTGCTTACTACAGAGATCAAGATCAAGATCAAGATCTTCGAGTAC-----ACC 103
QY 21 GlyIySmetValSerIySerIleAlaThrIleuIleAspAspThrSerSerGluVal 40
    |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 104 TCTTAATCACTGGAAGAAAGAAATCGAGGTAATATCTTGAGGAGATATGATCAAGATCC 163
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnIySylleuGluVal 60
    |||  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::  ::
DB 164 CAGGATCAGCTGATGAATTTGACTTCAAGAGATCAAGAACTTAAGATGAATCCAGAA 223

```

```
Qy 61 LysIleuYsAsnIleuIleuThrValIleuYsIleuAlaIleuIleuTyraArgAsnAngln 80
Db 224 AGAGAGAGAGAACTG-----GAGTTGAATGTGTATCAGAGACAACTG 268
Qy 81 PheAsnGlnAspGlnIleuAlaIleuMetGlnLys-----PheLysLysLysValHisGln 98
Db 269 AATGTGCAATTAAATTCTTCTCTCGAGAGAGATGAATCTAATAAAGAGCTTCATGCT 328
Qy 99 LeuAlaMetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValIleu 118
Db 329 ATTCAAAAGGAATACAGTGAATTCAGTCC-----AATCTGATGAGAAAGTGCT 379
Qy 119 Ser-----ArgLeuLeuAsnGlnCysArgGlnMetLeuHisGlnIleIleGlnArgHis 136
Db 380 TCAGATCGAAGCTGTGGGGGACCAAGAGAGAGACTCACCGCTTGAGGACCAACTT 439
Qy 137 LeuThrAlaLysSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGln 156
Db 440 GGCACTGCTTAAGTGAAAGCAAGTAAATGAAGTG----- 475
Qy 157 PheLeuAlaAlaLeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCys 176
Db 476 CTAATTGCTGATCTGACTCGAAGAAAAGAAAACCTTAGAAGAAATGATGATGCTGAGCTG 535
Qy 177 AspGlyTLeuAsnLysMetLeuAspGln 185
Db 536 GACAAATGTAAACAAGTTAAAGCAAGAG 562
```

Search completed: July 28, 2005, 19:22:51
Job time : 563 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 19:01:22 ; Search time 622 Seconds

(without alignments)
1956.275 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962

Sequence: 1 MARDVRSKRLAVQAQKIL.....KPHLQKLCDCGINKMLDEBNT 188

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Command line parameters:
-MODELS=frame+ p2n.model -DEV=xjh
-O=/cg2_1/USPTO.spool.h/US10627571/runat.27072005.154722.27521/app.query.fasta_1.327
-DB=Published Applications NA -QFMT=fastap -SUPERX=rnpb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsuns62
-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR_SCORE=spct -THR_MAX=100
-THR_MIN=0 -ALIGN=45 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10627571 @CGN 1 1 480 @runat.27072005.154722.27521
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WMT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1:	/cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2:	/cg2_6/ptodata/2/pubpna/PTC_NEW_PUB.seq:*
3:	/cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4:	/cg2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5:	/cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6:	/cg2_6/ptodata/2/pubpna/PTCUS_PUBCOMB.seq:*
7:	/cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8:	/cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9:	/cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10:	/cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11:	/cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
12:	/cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
13:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
14:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
15:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
16:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
17:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
18:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
19:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
20:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
21:	/cg2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
22:	/cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
23:	/cg2_6/ptodata/2/pubpna/US11_PUBCOMB.seq:*
24:	/cg2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25:	/cg2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26:	/cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

Score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	954	99.2	1892	22 US-10-491-545A-48	Sequence 48, Appl
2	954	99.2	1915	18 US-10-627-571-1	Sequence 1, Appl1
3	950	98.8	1814	19 US-10-755-889-215	Sequence 215, Appl1
4	950	98.8	2003	19 US-10-755-889-217	Sequence 217, Appl
5	950	98.8	2034	13 US-10-087-192-497	Sequence 497, Appl
6	950	98.8	2081	19 US-10-755-889-39	Sequence 39, Appl1
7	950	98.8	58723	13 US-10-087-192-496	Sequence 496, Appl
8	948	98.5	1943	14 US-10-097-065-98	Sequence 98, Appl1
9	920	95.6	2087	13 US-10-087-192-494	Sequence 494, Appl
10	920	95.6	62231	13 US-10-087-192-493	Sequence 493, Appl
11	914	95.0	544	16 US-10-029-386-22914	Sequence 22914, A
12	914	95.0	544	16 US-10-029-386-22914	Sequence 25411, A
13	610	63.4	527	16 US-10-094-466-43	Sequence 43, Appl1
14	593	61.6	619	17 US-10-094-466-41	Sequence 41, Appl1
15	593	61.6	645	17 US-10-094-466-41	Sequence 70, Appl1
16	593	61.6	2108	18 US-10-416-314-70	Sequence 5, Appl1
17	591	61.4	1188	9 US-09-816-828-5	Sequence 2116, Ap
18	566	58.8	1389	21 US-10-764-420-2116	Sequence 81, Appl1
19	527	54.8	1113	9 US-09-822-849A-81	Sequence 15, Appl1
20	527	54.8	1165	21 US-10-363-374-15	Sequence 537, Appl
21	527	54.8	1175	18 US-10-302-172-537	Sequence 59, Appl1
22	527	54.8	1268	21 US-10-959-539-59	Sequence 3207, Ap
23	504	52.4	287	9 US-09-796-692-3207	Sequence 3207, Ap
24	504	52.4	287	14 US-10-040-862-3207	Sequence 3207, Ap
25	504	52.4	287	17 US-10-057-475B-3207	Sequence 3207, Ap
26	504	52.4	287	17 US-10-154-884B-3207	Sequence 3207, Ap
27	504	52.4	287	19 US-10-764-324-3207	Sequence 2801, Ap
28	499	51.9	287	9 US-09-796-692-2801	Sequence 2801, Ap
29	499	51.9	287	14 US-10-040-862-2801	Sequence 2801, Ap
30	499	51.9	287	17 US-10-057-475B-2801	Sequence 2801, Ap
31	499	51.9	287	17 US-10-154-884B-2801	Sequence 2801, Ap
32	499	51.9	287	19 US-10-764-324-2801	Sequence 5724, Ap
33	484	50.3	288	9 US-09-796-692-5724	Sequence 5724, Ap
34	484	50.3	288	14 US-10-040-862-5724	Sequence 5724, Ap
35	484	50.3	288	17 US-10-057-475B-5724	Sequence 5724, Ap
36	484	50.3	288	17 US-10-154-884B-5724	Sequence 5724, Ap
37	484	50.3	422	10 US-09-918-999-35521	Sequence 35521, A
38	324	33.7	422	10 US-10-029-386-9211	Sequence 9211, A
39	304	31.6	565	22 US-10-972-079-232	Sequence 232, Appl
40	189	19.6	543	17 US-10-191-803-330	Sequence 11708, A
41	155	16.1	433	17 US-10-029-386-11708	Sequence 4917, Ap
42	148	15.4	538	16 US-09-796-692-4917	Sequence 4917, Ap
43	109	11.3	538	14 US-10-040-862-4917	Sequence 4917, Ap
44	109	11.3	538	17 US-10-057-475B-4917	Sequence 4917, Ap
45	109	11.3	538	17 US-10-154-884B-4917	Sequence 4917, Ap
46	109	11.3	538	17 US-10-764-324-4917	Sequence 6328, Ap
47	109	11.3	538	19 US-10-764-324-4917	Sequence 6328, Ap
48	95	9.9	638	9 US-09-974-300-6328	Sequence 167, App
49	93.5	9.7	2343	17 US-10-282-122A-10827	Sequence 21754, A
50	92.5	9.6	2501	14 US-10-198-846-13958	Sequence 13958, A
51	91	9.5	1305	21 US-10-831-070-279	Sequence 279, App
52	90.5	9.4	4256	9 US-09-745-763-167	Sequence 167, App
53	90.5	9.4	4256	10 US-09-814-355-21754	Sequence 445, App
54	90.5	9.4	4428	17 US-10-295-021-445	Sequence 133, App
55	90	9.4	1887	16 US-10-349-680-133	Sequence 8, Appl1
56	89.5	9.3	1142	13 US-10-082-018-8	Sequence 8, Appl1
57	89.5	9.3	1142	21 US-10-949-978-8	Sequence 623, Appl
58	88.5	9.2	1700	9 US-09-954-531-623	Sequence 1081, Ap
59	88.5	9.2	1700	17 US-10-305-720-1081	Sequence 898, App
60	88.5	9.2	1700	18 US-10-641-643-898	Sequence 1690, Ap
61	88.5	9.2	1700	21 US-10-843-641A-1690	Sequence 9360, Ap
62	88.5	9.1	1109	19 US-10-767-701-9360	Sequence 46, Appl
63	88	9.1	1638	9 US-09-976-740-46	Sequence 46, Appl
64	88	9.1	1638	13 US-10-023-529-46	Sequence 46, Appl
65	88	9.1	1638	13 US-10-023-529-46	Sequence 46, Appl

Percent Similarity:	99.47%	Conservative:	0
Best Local Similarity:	99.47%	Mismatches:	1
Query Match:	99.17%	Indels:	0
DB:	22	Gaps:	0
US-10-627-571-2 (1-188) x US-10-491-545A-48 (1-1892)			
Qy	1 MetLathrArpValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIlySerIleLeu 20		
Db	98 ATGGCCACAGATGCTTTAATTCAAAACTGGCGCTTAGCGCCAAAGAAGATCTTG 157		
Qy	21 GlySerMetValSerIysSerIleAlaThrThrIleuIleAspPthSerSerGluVal 40		
Db	158 GGTAAATGCTGTCCAATTCATCGCCACCACTTAAATAGACACACAAGATGAGGCG 217		
Qy	41 LeuAspGluLeuIYrArgValIThrArgGluIYrThrGlnAsnIlySylGluAlaGluIys 60		
Db	218 CTGATGATGCTCTACAGTGACAGGACAGATACCCAAAAACAAGAGAGCAAGAAG 277		
Qy	61 LysIleIysAsnLeuIleIysThrValIleIysIleuAlaIleLeuIYrArgAsnAsnGln 80		
Db	278 ATCATCAAGAACCTCATCAACAGCATCATCAAGCTGGCATCTCTTAATAGGAATATACG 337		
Qy	81 PheAsnGlnAspGluLeuAlaLeuMetGluIysPheIlySylSylValIHisGlnLeuAla 100		
Db	338 TTTAATCAAGATGAGCTAGCATGTATGAGAAATTTAAGAGAAAGTTCATCAGCTTCT 397		
Qy	101 MetThrValIleSerPheHisGlnValAspIYrThrPheAspArgAsnValIleuSerArg 120		
Db	398 ATGACCGTGCTCAATTCATCAGTGGATTAATACCTTTGACCGGAATGTGTATCCAGG 457		
Qy	121 LeuLeuAsnGluCybArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140		
Db	458 CTGTTAAATGATCAGAGAGATGCTGCACCAAAATCATTCAGCGCACCTCATCTCCAAAG 517		

```

RESULT 2
US-10-627-571-1
; Sequence 1, Application US/10627571
; Publication No. US20040082771A1
; GENERAL INFORMATION:
;   APPLICANT: KASID, Usha N.
;   APPLICANT: KUMAR, Deepak
;   APPLICANT: GOKHALE, Pratulla
;   APPLICANT: AHMAD, Imran
; TITLE OF INVENTION: ANTI-APOTOPIC GENE SCC-S2 AND DIAGNOSTIC AND THERAPEUTIC USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 223316
; CURRENT APPLICATION NUMBER: US/10/627,571
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: US 60/264,062
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/US02/02212
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature

```

OTHER INFORMATION: SCC-S2
US-10-627-571-1

Alignment Scores:
Pred. No.: 5,85e-97 Length: 1915
Score: 954.00 Matches: 187
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.17% Indels: 0
DB: 18 Gaps: 0

US-10-627-571-2 (1-188) x US-10-627-571-1 (1-1915)

```

QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAglInLySylLeu 20
DB 134 ATGGCCACAGATGCTTTAATTCGAAAACCTGGCCCTTACAGGCAAAAAGATCTTG 193
QY 21 GlyMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 194 GGTAAATGGTGTCCAAATCCATGCGCACCTTAATAGCAGACAGTAGTAGGCTG 253
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySylGluAglLys 60
DB 254 CTGGATGAGCTCTACAGAGTGACCGAGGAGTACCCAAACAAAGAGGAGGAGAG 313
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAsnGln 80
DB 314 ATCATCAAGAACCTCATCAAGACAGTCATCAGTCCATCTTTATAGGAATTAATCAG 373
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAla 100
DB 374 TTATATCAAGATGAGTATGATGAGGAAATTTAAGAGAAAGTTCATCAGCTTGTCT 433
QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
DB 434 ATGACGGTGTGCTGATTTCCATCAGTGGATTAATCCTTTGACCGGAATGTATTACAGG 493
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 494 CTGTAAATCAATGACAGAGAGATGCTCCACCAATCATTCAGCGCCACCTCAGTCCAG 553
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
DB 554 TCACATGAGACGGGTAAATAGTCTTGATCATTTTCAATGTGTGAATTTTGGCTGCC 613
QY 161 LeuTyAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 614 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGATGTATCAAC 673
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 674 AAAATGTTGGATGAGAGAACATTA 697

```

RESULT 3
US-10-755-889-215
; Sequence 215, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 215
; LENGTH: 1814
; TYPE: DNA

ORGANISM: Homo sapiens
US-10-755-889-215

Alignment Scores:
Pred. No.: 1.52e-96 Length: 1814
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 19 Gaps: 0

US-10-627-571-2 (1-188) x US-10-755-889-215 (1-1814)

```

QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAglInLySylLeu 20
DB 174 GTGGCCACAGATGCTTTAATTCGAAAACCTGGCCCTTACAGGCAAAAAGATCTTG 223
QY 21 GlyMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 234 GGTAAATGGTGTCCAAATCCATGCGCACCTTAATAGCAGACAGTAGTAGGCTG 293
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySylGluAglLys 60
DB 294 CTGGAGAGGCTCTACAGAGTGACCGAGGAGTACCCAAACAAAGAGGAGGAGAG 353
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAsnGln 80
DB 354 ATCATCAAGAACCTCATCAAGACAGTCATCAGTCCATCTTTATAGGAATTAATCAG 413
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAla 100
DB 414 TTATATCAAGATGAGTATGATGAGGAAATTTAAGAGAAAGTTCATCAGCTTGTCT 473
QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
DB 474 ATGACGGTGTGCTGATTTCCATCAGTGGATTAATCCTTTGACCGGAATGTATTACAGG 533
QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 534 CTGTAAATCAATGACAGAGAGATGCTCCACCAATCATTCAGCGCCACCTCAGTCCAG 593
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
DB 594 TCACATGAGACGGGTAAATAGTCTTGATCATTTTCAATGTGTGAATTTTGGCTGCC 653
QY 161 LeuTyAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 654 TTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGATGTATCAAC 713
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 714 AAAATGTTGGATGAGAGAACATTA 737

```

RESULT 4
US-10-755-889-217
; Sequence 217, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 217
; LENGTH: 2003
; TYPE: DNA


```
/ ORGANISM: Homo sapiens
US-10-755-889-217

Alignment Scores:
Pred. No.: 1,77e-96      Length: 2003
Score: 950.00           Matches: 186
Percent Similarity: 99.47%  Conservative: 1
Best Local Similarity: 98.94%  Mismatches: 1
Query Match: 98.75%      Indels: 0
DB: 19                  Gaps: 0

US-10-627-571-2 (1-188) x US-10-755-889-217 (1-2003)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 204 GTGGCCACAGATGCTTTTAATTCCAAAAACCTGGCCGTTCAGCACAAAGAAATCTTG 263

QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 264 GGTAATAATGCTGCCAATTCATCGCCACCACTTATATAGACACAAAGTAGTGAGTG 323

QY 41 LeuAspGluLeuTyrrArgValThrArgLutTyrrThrGlnAsnLysLysGluAlaGluLys 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 324 CTGGATGAGCTCTTACAGAGTACAGGAGTACACCAAAACAAAGAGGAGGAGGAAG 383

QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnAsnGln 80
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 384 ATCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGCCATCTTTATAGAAATATCAG 443

QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 444 TTTAATCAAGATGAGCTGATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCT 503

QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 504 ATGACCGTGTCAGTTCATCAAGGAGTATTAACCTTTGACCGGATGTTATCCAGG 563

QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 564 CTGTTAAATGAATGACGAGAGAGTCTGCACCAATATATTCAGGCCACCTCAGCTCCAG 623

QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 624 TCACATGACGCGGTTAATATGTTGATTCATTTTTCAGATTTGAAATTTTGGCTGCC 683

QY 161 LeuTyrrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuAspGluLysIleAsn 180
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 684 TTGTATTAATCCTTTGGGAAATTTTAAACCCCACTTACAAAACTATGTATGTATCAAC 743

QY 181 LysMetLeuAspGluGluAsnIle 188
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 744 AAAATGTTGATGAAGAGAAACATA 767

RESULT 5
US-10-087-192-497
/ Sequence 497, Application US/10087192
/ Publication No. US20020182586A1
/ GENERAL INFORMATION:
/ APPLICANT: Morris, David W.
/ APPLICANT: Engelhard, Eric K.
/ TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
/ FILE REFERENCE: 529452000122
/ CURRENT APPLICATION NUMBER: US/10/087,192
/ CURRENT FILING DATE: 2002-03-01
/ PRIOR APPLICATION NUMBER: US 09/747,377
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: US 09/798,586
/ PRIOR FILING DATE: 2001-03-02
/ NUMBER OF SEQ ID NOS: 2059
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 497
/ LENGTH: 2034
```

```
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-087-192-497

Alignment Scores:
Pred. No.: 1,81e-96      Length: 2034
Score: 950.00           Matches: 186
Percent Similarity: 99.47%  Conservative: 1
Best Local Similarity: 98.94%  Mismatches: 1
Query Match: 98.75%      Indels: 0
DB: 13                  Gaps: 0

US-10-627-571-2 (1-188) x US-10-087-192-497 (1-2034)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 248 GTGGCCACAGATGCTTTTAATTCCAAAAACCTGGCCGTTCAGCACAAAGAAATCTTG 307

QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 308 GGTAATAATGCTGCCAATTCATCGCCACCACTTATATAGACACAAAGTAGTGAGTG 367

QY 41 LeuAspGluLeuTyrrArgValThrArgLutTyrrThrGlnAsnLysLysGluAlaGluLys 60
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 368 CTGGATGAGCTCTTACAGAGTACAGGAGTACACCAAAACAAAGAGGAGGAGGAAG 427

QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrrArgAsnAsnGln 80
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 428 ATCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGCCATCTTTATAGAAATATCAG 487

QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 488 TTTAATCAAGATGAGCTGATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCT 547

QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLeuSerArg 120
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 548 ATGACCGTGTCAGTTCATCAAGGAGTATTAACCTTTGACCGGAAATGTTATCCAGG 607

QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 608 CTGTTAAATGAATGACGAGAGAGTCTGCACCAATATATTCAGGCCACCTCAGCTCCAG 667

QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 668 TCACATGACGCGGTTAATATGTTGATTCATTTTTCAGATTTGAAATTTTGGCTGCC 727

QY 161 LeuTyrrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGluLysIleAsn 180
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 728 TTGTATTAATCCTTTGGGAAATTTTAAACCCCACTTACAAAACTATGTATGTATCAAC 787

QY 181 LysMetLeuAspGluGluAsnIle 188
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 788 AAAATGTTGATGAAGAGAAACATA 811

RESULT 6
US-10-755-889-39
/ Sequence 39, Application US/10755889
/ Publication No. US20040171823A1
/ GENERAL INFORMATION:
/ APPLICANT: Bristol-Myers Squibb Company
/ APPLICANT: Bristol-Myers Squibb Company
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
/ FILE REFERENCE: D0284 NP
/ CURRENT APPLICATION NUMBER: US/10/755,889
/ CURRENT FILING DATE: 2004-01-13
/ PRIOR APPLICATION NUMBER: U.S. 60/440,068
/ PRIOR FILING DATE: 2003-01-14
/ PRIOR APPLICATION NUMBER: U.S. 60/469,757
/ PRIOR FILING DATE: 2003-05-12
/ NUMBER OF SEQ ID NOS: 823
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 39
/ LENGTH: 2081
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-755-889-39

Alignment Scores:
Pred. No.: 1,876-96 Length: 2081
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 19 Gaps: 0

US-10-627-571-2 (1-188) x US-10-755-889-39 (1-2081)

QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAglInuLySylleu 20
DB GTGGCCACAGATGCTTTTAATTCCAAAACCTGGCGCTCAGGACCAAAAGAGATCTTG 326
QY 21 GlyLyMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB GTTAAATGGTGTCCAAATCCATCGCCACACCTTAATAGACACCAAGTAGTGAGTGG 386
QY 41 LeuAspGluLeuTyraArgValThrArgGluTyThrGlnAsnLySylleuAglInuLyS 60
DB CTGGATGAGCTTACAGAGTGAACAGGAGTACACCAAAACAAAGAGAGGAGAGAG 446
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnAgn 80
DB ATCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGAAATATCAG 506
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLySphelysValHisGlnLeuAla 100
DB TTTTAATCAAGATAGAGTACGATTTGATGAGAAATTTAAGAAAGTTCATCAGCTTGC 566
QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValleuSerArg 120
DB ATGACCGTGTGCTGATTTCCATCAGTGAGTATTAACCTTTGACCGAATGTGTTATCCAG 626
QY 121 LeuLeuAsnGluCyAsnArgLumeLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB CTGTTAAATGAATGACAGAGAGATGCTGCACCAATATATTACAGGCCACTGCACAG 686
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCySgluPheLeuAlaLys 160
DB TCACATGAGACGGGTATATATGTTGATTCATTTTTCAGATTGTGAATTTTGGCTGCC 746
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
DB TTGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAAC 806
QY 181 LysMetLeuAspGluGlnAsnIle 188
DB 807 AAAATGTTGATGAAGAAACATA 830

RESULT 7
US-10-087-192-496
; Sequence 496, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 496

```

```

; LENGTH: 58723
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-496

Alignment Scores:
Pred. No.: 2,866-94 Length: 58723
Score: 950.00 Matches: 186
Percent Similarity: 99.47% Conservative: 1
Best Local Similarity: 98.94% Mismatches: 1
Query Match: 98.75% Indels: 0
DB: 13 Gaps: 0

US-10-627-571-2 (1-188) x US-10-087-192-496 (1-58723)

QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAglInuLySylleu 20
DB GTGGCCACAGATGCTTTTAATTCCAAAACCTGGCGCTCAGGACCAAAAGAGATCTTG 4696
QY 21 GlyLyMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB GTTAAATGGTGTCCAAATCCATCGCCACACCTTAATAGACACCAAGTAGTGAGTGG 47056
QY 41 LeuAspGluLeuTyraArgValThrArgGluTyThrGlnAsnLySylleuAglInuLyS 60
DB CTGGATGAGCTTACAGAGTGAACAGGAGTACACCAAAACAAAGAGAGGAGAGAG 47116
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraArgAsnAgn 80
DB ATCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGAAATATCAG 47176
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLySphelysValHisGlnLeuAla 100
DB TTTTAATCAAGATAGAGTACGATTTGATGAGAAATTTAAGAAAGTTCATCAGCTTGC 47236
QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValleuSerArg 120
DB ATGACCGTGTGCTGATTTCCATCAGTGAGTATTAACCTTTGACCGAATGTGTTATCCAG 47296
QY 121 LeuLeuAsnGluCyAsnArgLumeLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB CTGTTAAATGAATGACAGAGAGATGCTGCACCAATATATTACAGGCCACTGCACAG 47356
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCySgluPheLeuAlaLys 160
DB TCACATGAGACGGGTATATATGTTGATTCATTTTTCAGATTGTGAATTTTGGCTGCC 47416
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
DB TTGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGTGATGATCAAC 47476
QY 181 LysMetLeuAspGluGlnAsnIle 188
DB 47477 AAAATGTTGATGAAGAAACATA 47500

RESULT 8
US-10-097-065-98
; Sequence 98, Application US/10097065
; Publication No. US20030055236A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/097,065
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18

```

```

; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,053
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,064
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,054
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,008
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,365
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1926)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1928)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1934)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-097-065-98

Alignment Scores:
Pred. No.: 2,84e-96 Length: 1943
Score: 948.00 Matches: 186
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 2
Query Match: 98.54% Indels: 0
DB: 14 Gaps: 0

US-10-627-571-2 (1-188) x US-10-097-065-98 (1-1943)
QY 1 MetAlaThAspValPheAsnSerLyAsnLeuAlaValAlaGlnIleuValysIleLeu 20
DB 120 ATGGCCACGATGTCCTTAATCCAAAACCTGGCGTTCANGCACAAAAGAAATCTTG 179
QY 21 GlyLysMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 180 GGTAAATGTCCTCAATCCATCCGACCACTTAATGACACCAAGTACGAGTG 239
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLysLysGluAlaGluLys 60
DB 240 CTGGATGAGCTCTACAGAGTGACACAGGAGTACACCAAAACAAGAGGAGGAGAG 299
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAngin 80
DB 300 ATCATCAAAACCTCATCAAGACAGCATCAAGCTGCGCATCTTTATGGAATATCAG 359
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla 100
DB 360 TTTATCAAGATGAGTACATTGATGAGGAATTTTAAGAAAGAAATTCATCAGCTTGC 419
QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
|||||
```

```

DB 420 ATGACCGTCGTCAGTTCCATCAGGTGATATACCTTTGACCGGAATGTATCCAGG 479
QY 121 LeuIleuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 480 CTGTTAAATGAATGACAGAGAGTGTGACCAAAATCATTCACGCGCACCCTCCTCCAG 539
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
DB 540 TCACATGACGCGGTATTAATATGTGTGATCATTTTTCAGATTGGAATTTTGGCTGCC 599
QY 161 LeuTyAsnProPheGlyAsnLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 600 TTGTATTAATCCTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATGTCAAC 659
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 660 AAAATGTTGATGAAGAAGAACATA 683
|||||
```

RESULT 9

```

US-10-372-876-98
; Sequence 98, Application US/10372876
; Publication No. US20030204071A1
; GENERAL INFORMATION:
; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021p1
; CURRENT APPLICATION NUMBER: US/10/372,876
; CURRENT FILING DATE: 2003-02-26
; PRIOR APPLICATION NUMBER: 09/334,595
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1926)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1928)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1934)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-372-876-98
```

Alignment Scores:

Pred. No.: 2,846-96 Length: 1943
Score: 948.00 Matches: 186
Percent Similarity: 98.94% Conservative: 0
Best Local Similarity: 98.94% Mismatches: 2
Query Match: 98.54% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-372-876-98 (1-1943)

```
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnIleGlnLysIleLeu 20
DB 120 ATGGCCACAGATGCTTTAATTCACAAAACCTGGCCCTTANGCACAAAAGAAATCTTG 179
QY 21 GlyLysMetValSerLysSerIleAlaThrThreuleIleAspThrSerSerGluVal 40
DB 180 GGTAATAATGGTGTCCAAATCCATCGCCACCACTTAATAGCAGACAAAGATGAGAGTG 239
QY 41 LeuAspGluLeuTyraGValThrArgGluTyThrGlnAsnLysGluAlaGluLys 60
DB 240 CTGATAGAGCTCTTACAGAGTACCGAGGAGTACCCCAAAACAAAGAGAGGAGAGAG 299
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraGAsnGln 80
DB 300 ATCATCAAGAACCTCATCAAGACGATCATCAGCTGGCCATCTTTATAGAAATATCAG 359
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla 100
DB 360 TTTAATCAAGATGAGTATGATGATGAGAAATTTAAGAAAGAAAGTTCAATCAGCTGCT 419
QY 101 MetThrValIleSerPheIleGlnValAspTyThrPheAspArgAsnValIleSerArg 120
DB 420 ATGACCGTGTCTGATTTCCATCAGGTGATTTAATCTTTAATCGGAATGTGTTATCCAG 479
QY 121 LeuLeuAsnGluCysArgGluMetLeuIleGlnIleGlnArgHisLeuThrAlaLys 140
DB 480 CTGTTAATGAATGACGAGAGAGATGCTGCACCAATCTTTCAGCCCACTCATCTGCAG 539
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 540 TCACATGACGAGGCTTATATATGTTGATCATTTTTCAGATGTGAAATTTTGGCTGCC 599
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 600 TTGTATATATCTTTTGGAAATTTTAAACCCCACTTACAAAACTATGTGATGTATCAAC 659
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 660 AAAATGTGATGAAGAGAAACATA 683
```

RESULT 10

US-10-087-192-494
; Sequence 494, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 494
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-494

Alignment Scores:

Pred. No.: 4,576-93 Length: 2087
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
DB: 13 Gaps: 0

US-10-627-571-2 (1-188) x US-10-087-192-494 (1-2087)

```
QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnIleGlnLysIleLeu 20
DB 507 GTGGCTACAGATGCTTTCAATTCACAAAACCTGGCCCTTANGCACAAAAGAAATCTTG 566
QY 21 GlyLysMetValSerLysSerIleAlaThrThreuleIleAspThrSerSerGluVal 40
DB 567 GGCAAAATGGTATCCAAATCCATCGCCACCACTTATGACGACCAACGAGAGAGTG 626
QY 41 LeuAspGluLeuTyraGValThrArgGluTyThrGlnAsnLysGluAlaGluLys 60
DB 627 CTAGATGAGCTCTTACAGGAGTACCAAGAGTATACCCCAAGAAAGAGAGGAGAGG 686
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyraGAsnGln 80
DB 687 GTCATCAAGAACCTCATCAAGACGATCATCAGCTGGCCATCTTCCACAGAAACAAATCAG 746
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValIleGlnLeuAla 100
DB 747 TTCAATCAAGACGATGAGGCTCATGAGAAATTTCAAGAAAGAGGAGGAGGAGGAGG 806
QY 101 MetThrValIleSerPheIleGlnValAspTyThrPheAspArgAsnValIleSerArg 120
DB 807 ATGACCGTGTCTGATTTCCATCAGGTGATTTAATCTTTAATCGGAATGTGTTATCCAG 866
QY 121 LeuLeuAsnGluCysArgGluMetLeuIleGlnIleGlnArgHisLeuThrAlaLys 140
DB 867 CTGCTAATCAAGAGGCTTATATATGCTTTGACCAATTTTTCAGATGTGATTTTGGCTGCC 926
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 927 TCTCAAGAGAGGCTTATATATGCTTTGACCAATTTTTCAGATGTGATTTTGGCTGCC 986
QY 161 LeuTyraAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
DB 987 TTGTATATATCTTTTGGAAATTTTAAACCCCACTTACAAAACTTATGTGATGTATCAAC 1046
QY 181 LysMetLeuAspGluGluAsnIle 188
DB 1047 AAAATGTGATGAAGAGAAACATA 1070
```

RESULT 11

US-10-087-192-493
; Sequence 493, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 62231
; TYPE: DNA
; ORGANISM: Mus musculus

```

; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(62231)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-493

Alignment Scores:
Pred. No.: 7,56e-91 Length: 62231
Score: 920.00 Matches: 177
Percent Similarity: 98.94% Conservative: 9
Best Local Similarity: 94.15% Mismatches: 2
Query Match: 95.63% Indels: 0
Gaps: 0

US-10-627-571-2 (1-188) x US-10-087-192-493 (1-62231)

QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLySylleu 20
Db 50651 GTGGCTACAGATGCTTTCATTCCAAAAACCTGGCGTTTCAGGACCAAAAAGATCTTG
QY 21 GlyMetValSerLySerLeuAlaThrThrLeuLeuAspAspThrSerSerGluVal 40
Db 50711 GGGAAATGGATCCAAATCCATCGCCACCGCTATGACACACACGACGAGGAGG
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySylleuAlaGln 60
Db 50771 CTAGATGAGCTGTACAGGGTGACCAAGAGTACACCCAGAACAAAGAGGCGGAGAGG
QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAngln 80
Db 50831 GTCATCAAAACCTCATCAAGAGGTCATCAAGCTGCGCTCTCCACAGGAACAATCAG
QY 81 PheAsnGluAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
Db 50891 TTCATCAAGACAGAGCTGGCGCTCATGAGAGATTCAGAGAGAGGACCGAGCTTGC
QY 101 MetThrValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
Db 50951 ATGAGCGGTGCTGAGCTTCCACCGAGTACAGCTTGACCGCATGTGCTGCAGG
QY 121 LeuLeuAsnGluCyAsnArgLueMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
Db 51011 CTGCTGAAAGAGGCGGAGGCTCTTACACGAGATATTTCAGCGCACCTTACCGCAAG
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
Db 51071 TCTCAGCGAGCGGCTTAATATGCTTTGACCATTTTTCAGATTGTGATTTTGGCTGCC
QY 161 LeuTyArgnProPheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
Db 51131 TTGTACATCCCTTTGGAAAGTTTAACCTTACAGAACTTTGGCAGCGGCATCAAC
QY 181 LysMetLeuAspGluGluAsnIle 188
Db 51191 AAAATGTGGATGAGAGAAACATA 51214

RESULT 12
US-10-029-386-22914
; Sequence 22914, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOICA-X-2
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22914
; LENGTH: 544

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC035144.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EST_HUMAN HIT: AL549492.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P31390, EVALUE 2.60e-01
; OTHER INFORMATION: NT HIT: G17657123, EVALUE 0.00e+00
US-10-029-386-22914

Alignment Scores:
Pred. No.: 2,87e-93 Length: 544
Score: 914.00 Matches: 179
Percent Similarity: 99.44% Conservative: 0
Best Local Similarity: 99.44% Mismatches: 1
Query Match: 95.01% Indels: 0
Gaps: 0

US-10-627-571-2 (1-188) x US-10-029-386-22914 (1-544)

QY 2 AlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLySylleuGly 21
Db 3 GGCACAGATGCTTTTATTCCAAAAACCTGGCGTTTCAGGACCAAAAAGATCTTGGGT
QY 22 LysMetValSerLySerLeuAlaThrThrLeuLeuAspAspThrSerSerGluValLeu 41
Db 63 AAAATGTGTCCAAATCCATCCACACCACTTAAATAGACACAAAGATGTAGGTGCTG
QY 42 AspGluLeuTyArgValThrArgGluTyThrGlnAsnLySylleuAlaGlnLysLys 61
Db 123 GATGAGCTGTACAGAGTGACAGGAGTACACCCAAACAAAGAGAGGACAGAGATGC
QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAnglnPhe 81
Db 183 ATCAAGAACTCATCAAGCACTCATCAAGCTGCGCATTTTATATGGAATATACGTT
QY 82 AsnGluAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
Db 243 AATCAAGATGAGCTGATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTCTAG
QY 102 ThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 121
Db 303 ACCGCTGTAGTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTATTCAGGCTG
QY 122 LeuAsnGluCyAsnArgLueMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
Db 363 TTAATGATATGACAGAGATGCTGCACCAATCATTTACAGCGGCACCTCAGTCCAAAGTCA
QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
Db 423 CATGACGGGCTTAATATGCTTTGATCATTTTTCAGATTGTGAAATTTTGGCTGCC
QY 162 TyraAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsnLys 181
Db 483 TATATCCCTTTGGAAATTTTAACCCCATTAACAAACATATGATGTATCAACAA 542

RESULT 13
US-10-029-386-25411
; Sequence 25411, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOICA-X-2
; CURRENT FILING DATE: 2001-12-20

```



```

0Y      62  ILuysAsnIuElIeYsThrVallIeYsIeuAlaIleuYrYrYAsnAsnIuPhe 81
Db      203 CTCAGAGAACTGGTCAAGGTGGCCCTGAAAGCTGGAGACTGCTGCGCGGAGACCACTG 262
0Y      82  AsnGlnAspGluIuEuAlaEuMeTcIuYsPheYsIySlySlyVAlHsGlnEuAlaMeT 107
Db      263 GGGCGGTGAGAGACTGGCGGTGCTGGCGCGCTTCCGCAACGAGCGCGCTGCTGCATG 322
0Y      102  ThrValIaSerPheHisGlnIValAspYrThrPheAspArgAsnValIeuSerArgIeu 121
Db      323 ACCGGCGCTCAGCTTCCACCAAGTGAAGCTTCACTTCCAGCCGCGGTCTGACCGCGCG 388
0Y      142  IleuAsnGluCysArGsluMeTleuHisGlnIleIeGlnArgHisIeUThrAlaIysSer 141
Db      383 CTGCTCGAGTGGCGCGGACCTGCTGCACCAAGCGCGTGCCTCCCACTGACCGCAAGTCC 444
0Y      142  HisGlyArYrYAlAsnAsnValIheAspHisPheSerAspYcGsluPheIuEuAlaIeu 161
Db      443 CACGGCGCGCATCAACCAACGCTGTTGGCGCACCTAGCGCGCATCGCACTTCTCGGCTGC 502
0Y      162  TyrAsnProPheGlyAsnPheIysProHisIeuGlnIlySleuYcAspAspGlyIleAsnIys 189
Db      503 TAGCGCCCCCGCCGAGCCCTTACCGCTTCCCACTGCGCAGAGATCTGAGAGGCTTGGCGCG 562
0Y      182  MetIeuAspGluGluAsnIle 188
Db      563 ATGCTGAGACGAGGCGACGCTC 583

RESULT 16
US-10-416-314-70
/ Sequence 70, Application US/10416314
/ Publication No. US20040082508A1
/ GENERAL INFORMATION:
/ APPLICANT: YUE, Henry
/ APPLICANT: YAO, Monique G.
/ APPLICANT: GANDHI, Ameena R.
/ APPLICANT: BAUGHN, Mariah R.
/ APPLICANT: SWARNAKAR, Anita
/ APPLICANT: CHAWLA, Narinder K.
/ APPLICANT: SANJANMALA, Madhusudan M.
/ APPLICANT: THORNTON, Michael B.
/ APPLICANT: ELLIOTT, Vicki S.
/ APPLICANT: LU, Yan
/ APPLICANT: GIETZEN, Kimberly J.
/ APPLICANT: BURFORD, Neil
/ APPLICANT: DING, Li
/ APPLICANT: HAFALIA, April J.A.
/ APPLICANT: TANG, Y. Tom
/ APPLICANT: BANDMAN, Olga
/ APPLICANT: WARREN, Bridget A.
/ APPLICANT: HONCHELL, Cynthia D.
/ APPLICANT: LU, Dying Aina M.
/ APPLICANT: THANGAVELU, Kavitha
/ APPLICANT: IEE, Sally
/ APPLICANT: XU, Yuming
/ APPLICANT: YANG, Junming
/ APPLICANT: TAL, Preeti G.
/ APPLICANT: TRAN, Bao
/ APPLICANT: ISON, Craig H.
/ APPLICANT: DUGGAN, Brendan M.
/ APPLICANT: KAREHT, Stephanie K.
/ TITLE OF INVENTION: SECRETED PROTEINS
/ FILE REFERENCE: PI-0287 USN
/ CURRENT APPLICATION NUMBER: US/10/416.314
/ CURRENT FILING DATE: 2003-05-08
/ PRIOR APPLICATION NUMBER: US 60/247,505
/ PRIOR FILING DATE: 2000-11-08
/ PRIOR APPLICATION NUMBER: US 60/249,642
/ PRIOR FILING DATE: 2000-11-09
/ PRIOR APPLICATION NUMBER: US 60/249,824
/ PRIOR FILING DATE: 2000-11-16
/ PRIOR APPLICATION NUMBER: US 60/252,824

```



```
; PRIOR FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: US 60/254,305
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: US 60/256,448
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PERL Program
; SEQ ID NO 70
; LENGTH: 2108
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 201309SCB1
US-10-416-314-70

Alignment Scores:
Pred. No.: 3,666-56          Length: 2108
Score: 593.00              Matches: 106
Percent Similarity: 79.14%   Conservative: 42
Best Local Similarity: 56.68% Mismatches: 39
Query Match: 61.64%         Indels: 0
DB: 18                      Gaps: 0

US-10-627-571-2 (1-188) x US-10-416-314-70 (1-2108)

QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeuGly 21
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 GCCATGGACACCTTCACACCAAGAGCGCTGCTCGACAGCGCGGAGAAAGAGCTTCGAGT 486

QY 22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 AAGATGGCGTCCAAAGGACGTGGCGCGCTGCTGATATACACCAAGCTGAGTGTCTG 546

QY 42 AspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGluLysLys 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 547 GATGAGCTGACCGCGCACAGGAGATTACAGCGCGCGCACAGCGCGCACAGGAGATG 606

QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPhe 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 607 CTCAGAGACCTGGTCAAGAGTGGCGCTGAGCTGAGCTGCTGGGAGACAGCTG 666

QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 667 GCGGCTGAGAGGTGGCGCTGCGCTGCGCGCTTCGCGCACCGGCGCGCTGCGCATG 726

QY 102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 727 ACCGCGCTCAGCTTCACACGATGACTTCACTTCGACCGGCGCGCTGCGCGCGG 786

QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 787 CTCCTGAGAGTCCGCGACCTGCTGACACAGGCGCGTGGTCCCACTGACCGGCAAGTCC 846

QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeu 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 847 CACGGCGGATCAACACAGTGTTCGCGCACCTAGCGACGACCTTCCTGCGTGGCGCTC 906

QY 162 TyrAsnProPheGlnAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLys 181
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 907 TACGGCCCGCCAGACCTACCGCTCCACCTGGACAGATGTGCGAGGCGCTGGCGCGG 966

QY 182 MetLeuAspGluGluAsnIle 188
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 967 ATGCTGGACGAGGCGACGCTC 987

RESULT 17
US-09-816-828-5
; Sequence 5, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
```

```
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yungting
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pc_FL_genes Version 2.0
; SEQ ID NO 5
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (174)..(731)
US-09-816-828-5

Alignment Scores:
Pred. No.: 2,66-56          Length: 1188
Score: 591.00              Matches: 106
Percent Similarity: 79.14%   Conservative: 42
Best Local Similarity: 56.68% Mismatches: 39
Query Match: 61.43%         Indels: 0
DB: 9                      Gaps: 0

US-10-627-571-2 (1-188) x US-09-816-828-5 (1-1188)

QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeuGly 21
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 GCCATGGACACCTTCACACCAAGAGCGCTGCTCGACAGCGCGGAGAAAGAGCTTCGAGT 230

QY 22 LysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 AAGATGGCGTCCAAAGGACGTGGCGCGCTGCTGATATACACCAAGCTGAGTGTCTG 290

QY 42 AspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGluLysLys 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 291 GATGAGCTGACCGCGCACAGGAGATTACAGCGCGCGCACAGGAGATGCGCGGAGAGATG 350

QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPhe 81
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 CTCAGAGACCTGGTCAAGAGTGGCGCTGAGCTGGACTGCTGCTGGGAGACAGCTG 410

QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 GCGGCTGAGAGACTGGCGCTGCTGCGCGCGCTTCGCCACCGGCGCGCTGCGCATG 470

QY 102 ThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 ACCGCGCTCAGCTTCACACGATGACTTCACTTCGACCGGCGCGCTGCGCGCTGCGG 530

QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 531 CTCCTGAGTGGCGGACCTGCTGACACAGGCGCGTGGCTCCCACTGACCGGCAAGTCC 590

QY 142 HisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeu 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 591 CACGGCGGATCAACACAGTGTTCGCGCACCTAGCGACGACCTTCCTGCGTGGCGCTC 650

QY 162 TyrAsnProPheGlnAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsnLys 181
```

```
Db      651 TACGGCCCGCCGAGCCCTACCGCTCCACCTCGCAGAGATCTCGCAGGGGCTTGCGGCCG 710
QY      182 MetLeuAspGluGluAsnIle 188
       |||
Db      711 ATGCTGACGAGGGCGAGCCTC 731

RESULT 18
US-10-764-420-2116
; Sequence 2116, Application US/10764420
; Publication No. US20050084872A1
; GENERAL INFORMATION:
; APPLICANT: Lum, Pek Yee
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764,420
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2116
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 80
; OTHER INFORMATION: n = A,T,C or G
US-10-764-420-2116

Alignment Scores:
Pred. No.: 2,18e-53 Length: 1389
Score: 566.00 Matches: 101
Percent Similarity: 79.46% Conservative: 46
Best Local Similarity: 54.59% Mismatches: 38
Query Match: 58.84% Indels: 0
Gaps: 0

US-10-627-571-2 (1-188) x US-10-764-420-2116 (1-1389)
QY      4 AspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeuGlyIysMet 23
       |||
Db      87 GACACCTTCAGCAGAAAGCCTGGCCCTGCAGGCCACAGAAAGTCTCCAGCAAGATG 146
QY      24 ValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGlu 43
       |||
Db      147 GCCTCCAAAGCCATGGTGGCTGTGTTGTGGACAATACCGACGTAGGCTTGGATGA 206
QY      44 LeuTyArgValThrArgGluTyThrGlnAsnIysGluAlaGluIysIleIys 63
       |||
Db      207 CTGTACCAAGCCACGAAAGAGTTCACGCGCACCGGAAGAGCAGACGAGGTAGTGA 266
QY      64 AsnLeuIleIysThrValIleIysLeuAlaIleLeuTyArgAsnAsnGlnPheAsnGln 83
       |||
Db      267 AACCTGTGAAGGTGGCTGTGAAGCTGGCTGTGTTCTGTGAGCGCAGACCGACTGCACAC 326
QY      84 AspGluLeuAlaLeuMetGluIysPheIysIysValHisGlnLeuAlaMetThrVal 103
       |||
Db      327 AATGAGCTGGCCCAAGTCGACGCGGTTCCGGGCGCGGTCCGCAACTGGCCATGACGCC 386
QY      104 ValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeuLeuAsn 123
       |||
Db      387 CTCAGCTTCACACAGGTAGACTTCACCTTGAACGGCGTGTGCGCACTGGGCTGCG 446
QY      124 GluGlyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIysSerHisGly 143
       |||
Db      447 GAGTCAGGAGCCTGTGTGACACGAGGCTATTGGCCCGCAGCTCACTGCAGTCCACGGC 506
```

```
QY      144 ArgValAsnAsnValPheAspHisPheSerAspGluPheLeuAlaLeuTyArgAsn 163
       |||
Db      507 CCATCATCATCATCTTCATCTTGCCTTGTGCAATGTGACTTCTGCGCGCGCTGACAC 566
QY      164 PropheGlyAsnPheIysProHisLeuGlnIysLeuCyAspGlyTLeaAsnIysMetLeu 183
       |||
Db      567 CCAGCAGAGCCCTACCGGAGCATCTGTGCGCATCTGTGATGCGCTCGGAGAGATGTTG 626
QY      184 AspGluGluAsnIle 188
       |||
Db      627 GAGAGGGGTGCATC 641

RESULT 19
US-09-822-849A-81
; Sequence 81, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakar
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822,849A
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 81
; LENGTH: 1113
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-81

Alignment Scores:
Pred. No.: 3,92e-49 Length: 1113
Score: 527.00 Matches: 99
Percent Similarity: 77.22% Conservative: 40
Best Local Similarity: 55.00% Mismatches: 39
Query Match: 54.78% Indels: 2
Gaps: 1

US-10-627-571-2 (1-188) x US-09-822-849A-81 (1-1113)
QY      6 PheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeuGlyIysMetValSer 25
       |||
Db      72 TTCAGTCCAAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 131
QY      26 IysSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGluLeuTyArg 45
       |||
Db      132 CGCTGTGCTCATCTCTTCATATGATGACAGACAGTAGAGGTGCTGATAGACTTAC 191
QY      46 ArgValThrArgGluTyThrGlnAsnIysGluAlaGluIysIleIysAsnLeu 65
       |||
Db      192 CGTGTGTCAAAGAGTACACGACGACCGCGCCGCGCCAGCGGTATCAAGGACCTG 251
QY      66 IleTyThrValIleIysLeuAlaIleLeuTyArgAsnAsnGlnPheAsnGlnAspGlu 85
       |||
Db      252 ATCAAGTGGCCATCAAGGTGGCTGTGCTGCAACCCCAATGCTCTTGGCCCCCAGTAG 311
QY      86 LeuAlaLeuMetGluIysPheIysIysValHisGlnLeuAlaMetThrValValSer 105
       |||
Db      312 CTGGCCCTGGCTACCGGCTTTGGCCAGAAAGCTGCGGACGGGTGCCATGACGCACTTAC 371
QY      106 PheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCy 125
       |||
```

Db 372 TTGTGATGAGTAGACTTACCTTCGAGCGTCTGTTGCGTGGCTGACCGAGTGC 431
 QY 126 ArgGlutMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIysSerHisGlyArgVal 145
 Db 432 CGGATGTGTGCTAGAGTTGGTGAACACCACTACCGCCCAAGTACATGCGCCGATC 491
 QY 146 AsnAsnValPheAspHisPheSerAspCyseGluPheLeuAlaAlaLeuTyrAsnProPhe 165
 Db 492 CGGCACGTGTTGATATCATCTTCTCTGACCCAGGTCGTCTACCGCCCTCTATGGGCTT 548
 QY 166 GlyAsnPheLysProHisLeuGlnIlyLeuCyseAspGlyTleAsnLysMetLeuAspGlu 185
 Db 549 ---GACTTCACCTACGACACTTGGCAAGATCTGTGACGAGCTACGAAAGCTGTAGAGAA 605

RESULT 20
 US-10-363-374-15
 ; Sequence 15, Application US/10363374
 ; Publication No. US20050048483A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Su, Eric
 ; APPLICANT: Wang, He
 ; TITLE OF INVENTION: NOVEL SECRETED PROTEINS AND THEIR USES
 ; FILE REFERENCE: X13020
 ; CURRENT APPLICATION NUMBER: US/10/363,374
 ; CURRENT FILING DATE: 2003-02-27
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 15
 ; LENGTH: 1165
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: CDS
 ; LOCATION: (100)..(651)
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: (100)..(180)
 US-10-363-374-15

Alignment Scores:
 Pred. No.: 4,26-49 Length: 1165
 Score: 527.00 Matches: 99
 Percent Similarity: 77.22% Conservative: 40
 Best Local Similarity: 55.00% Mismatches: 39
 Query Match: 54.78% Indels: 2
 Gaps: 1

US-10-627-571-2 (1-188) x US-10-363-374-15 (1-1165)

QY 6 PheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMetValSer 25
 Db 109 TTACAGCTCAAAAGAGCTCGGACCTGCAAGCAGAGAGAGAGCTACTAGTAGAGCGGGT 168
 QY 26 LysSerIleAlaThrThrIleuIleAspAspThrSerSerGluValLeuAspGluLeuTyr 45
 Db 169 CGCTCTGTGCTCATCTCTTCATAGATGAGACAAGAGAGTGAAGTGTAGTAGCTCTAC 228
 QY 46 ArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGlnLysLysIleLysAsnLeu 65
 Db 229 CGTGTGTCCAAAGAGTACACGACAGCGCGCCAGCGCCAGCGCGATCAAGACACTG 288
 QY 66 IleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnAspGlu 85
 Db 289 ATCAAAAGTGGCCATCAAGGTGCTGTGCTGACCGCATGTGCTTGTGCGCCAGTGTAG 348
 QY 86 LeuAlaLeuMetGluLysPheLysValHisGlnLeuAlaMetThrValValSer 105
 Db 349 CTGAGCTGTGCTACCGCGCTTTCGCAAGAGCTGGCGGAGGCTGATGAGGCACTTAGC 408
 QY 106 PheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125
 Db 409 TTGTGTGAGGTAGACTTCACTTCGAGAGCGTGTCTGTGCTGCGCTGTGACCGAGTGC 468

QY 126 ArgGlutMetLeuHisGlnIleIleGlnArgHisLeuThrAlaIysSerHisGlyArgVal 145
 Db 469 CGGATGTGTGCTAGAGTTGGTGAACACCACTACCGCCCAAGTACATGCGCCGATC 528
 QY 146 AsnAsnValPheAspHisPheSerAspCyseGluPheLeuAlaAlaLeuTyrAsnProPhe 165
 Db 529 CGGCACGTGTTGATATCATCTTCTCTGACCCAGGTCGTCTACCGCCCTCTATGGGCTT 585
 QY 166 GlyAsnPheLysProHisLeuGlnIlyLeuCyseAspGlyTleAsnLysMetLeuAspGlu 185
 Db 586 ---GACTTCACCTACGACACTTGGCAAGATCTGTGACGAGCTACGAAAGCTGTAGAGAA 642

RESULT 21
 US-10-302-172-537
 ; Sequence 337, Application US/10302172
 ; Publication No. US20040053250A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tang, Y. Tom
 ; APPLICANT: Xue, Aidong J.
 ; APPLICANT: Demanac, Radoje T.
 ; TITLE OF INVENTION: No. US20040053250A1el Arginine-rich Protein-like Nucleic Acids ar
 ; FILE REFERENCE: 803 1CNP
 ; CURRENT APPLICATION NUMBER: US/10/302,172
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 10/225,251
 ; PRIOR FILING DATE: 2002-08-20
 ; PRIOR APPLICATION NUMBER: PCT US02/05095
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 09/799,451
 ; PRIOR FILING DATE: 2001-03-05
 ; NUMBER OF SEQ ID NOS: 950
 ; SOFTWARE: pt_FL_genes Version 2.0
 ; SEQ ID NO 537
 ; LENGTH: 1175
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: CDS
 ; LOCATION: (113)..(664)
 US-10-302-172-537

Alignment Scores:
 Pred. No.: 4,26-49 Length: 1175
 Score: 527.00 Matches: 99
 Percent Similarity: 77.22% Conservative: 40
 Best Local Similarity: 55.00% Mismatches: 39
 Query Match: 54.78% Indels: 2
 Gaps: 1

US-10-627-571-2 (1-188) x US-10-302-172-537 (1-1175)

QY 6 PheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGlyLysMetValSer 25
 Db 122 TTACAGCTCAAAAGAGCTCGGACCTGCAAGCAGAGAGAGAGCTACTAGTAGAGCGGGT 181
 QY 26 LysSerIleAlaThrThrIleuIleAspAspThrSerSerGluValLeuAspGluLeuTyr 45
 Db 182 CGCTGTGTGCTCATCTCTTCATAGATGAGACAAGAGAGTGAAGTGTAGTAGCTCTAC 241
 QY 46 ArgValThrArgGluTyrThrGlnAsnLysLysGluAlaGlnLysLysIleLysAsnLeu 65
 Db 242 CGTGTGTCCAAAGAGTACACGACAGCGCGCCAGCGCCAGCGCGTGTCAAGACACTG 301
 QY 66 IleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGlnPheAsnGlnAspGlu 85
 Db 302 ATCAAAAGTGGCCATCAAGGTGCTGTGCTGACCGCATGTGCTTGTGCGCCAGTGTAG 361
 QY 86 LeuAlaLeuMetGluLysPheLysValHisGlnLeuAlaMetThrValValSer 105
 Db 362 CTGAGCTGTGCTACCGCGCTTTCGCAAGAGCTGGCGGAGGCTGATGAGGCACTTAGC 421
 QY 106 PheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeuLeuAsnGluCys 125

```

DB      422 TTTGGTAGAGTACCTTCACCTTCGAGGCTGCTGTTCTGGCTGGCTGGACCGAGTGC 481
QY      126 ArgGluMetLeuHisGlnIleIleGlnAArgHisLeuThrAlaLysSerHisGlyArgVal 145
DB      482 CGGAGATGTGCTGTAAAGTTGGTGAACACCACTTCACGCCCAAGTCACATGGCCGCTC 541
QY      146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeuTyraAsnPrope 165
DB      542 CGCAGAGTGTGATTAATCTTCTGTGACCGAGCTGTGCTACGCGCCCTCTATGGGCT--- 598
QY      166 GlyAsnPheLeuProHisLeuGlnIlyLeuCysAspGlyIleAsnLysMetLeuAspGlu 185
DB      599 ---GACTTCACCTACGACCTTGGCAAGATCTGTGACGAGCTCAGAACTGCTAAGACGAA 655

RESULT 22
US-10-959-539-59
/ Sequence 59, Application US/10959539
/ Publication No. US20050048623A1
/ GENERAL INFORMATION:
/ APPLICANT: INCYTE GENOMICS, INC.
/ APPLICANT: HILLMAN, Jennifer L.
/ APPLICANT: LAL, Preeti
/ APPLICANT: TANG, Y. Tom
/ APPLICANT: YUE, Henry
/ APPLICANT: AU-YOUNG, Janice
/ APPLICANT: BANDMAN, Olga
/ APPLICANT: AZIMZAI, Yalda
/ APPLICANT: YANG, Junming
/ APPLICANT: LU, Dying Aina M.
/ APPLICANT: BAUGHN, Mariah R.
/ APPLICANT: PATTERSON, Chandra
/ APPLICANT: SHAH, Purvi
/ TITLE OF INVENTION: CELL CYCLE AND PROLIFERATION PROTEINS
/ FILE REFERENCE: PF-0722 PCT
/ CURRENT APPLICATION NUMBER: US/10/959,539
/ PRIOR FILING DATE: 2004-10-06
/ PRIOR APPLICATION NUMBER: US/10/031,915
/ PRIOR FILING DATE: 2002-01-18
/ PRIOR APPLICATION NUMBER: 60/145,075; 60/153,129; 60/164,647
/ PRIOR FILING DATE: 1999-07-21; 1999-09-08; 1999-11-10
/ NUMBER OF SEQ ID NOS: 108
/ SOFTWARE: PERL Program
/ SEQ ID NO 59
/ LENGTH: 1268
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ OTHER INFORMATION: Incyte ID No: 1558289CB1
US-10-959-539-59

Alignment Scores:
Pred. No.: 4,77e-49 Length: 1268
Score: 527.00 Matches: 99
Percent Similarity: 77.22% Conservative: 40
Best Local Similarity: 55.00% Mismatches: 39
Query Match: 54,78% Indels: 2
DB: 21 Gaps: 1

US-10-627-571-2 (1-188) x US-10-959-539-59 (1-1268)
QY      6 PheAsnSerIlyAsnLeuAlaValGlnAlaGlnLysIleLeuGlnLysMetValSer 25
DB      112 TTCAGCTCAAGAGCTTGCACTGCAAGACGAGAAAGAGCTAGTAAGATGGCGGCT 171
QY      26 LysSerIleAlaThrThrLeuIleAspAspThrSerSerGlnValLeuAspGluLeuTy 45
DB      172 CGCTGTGGCTCATCTCTTCATAGATGAGACAAAGCAGTGGCTGTAGATGAGCTCTAC 231
QY      46 ArgValThrArgGluIlyThrGlnAsnLysLysGluAlaGlnLysLysIleLysAsnLeu 65
DB      232 CGTGTGTCAAGAGATACGCAACGCGGCGCCAGCGCCAGCGCTGATCAAGGACTG 291

```

```

QY      66 IleLysThrValIleLysLeuAlaIleLeuTyraAsnAsnGlnPheAsnGlnAspGlu 85
DB      292 ATCAAAAGTGCATCAAGATGGCTGTGCTGACCCGCAATGCTCTTGGCCCACTAG 351
QY      86 LeuAlaLeuMetGlnLysPheLeuLysLysValHisGlnLeuAlaMetThrValValSer 105
DB      352 CTGGCCCTGTGATACCGCTTGTCCAGAAAGCTGGCGGAGGGTGCATGACCGGCACTTAC 411
QY      106 PheHisGlnValAspTyThrPheAspArgAsnValLeuSerArgLeuLeuAsnLys 125
DB      412 TTTGGTAGAGTACCTTCACCTTCGAGGCTGCTGTTCTGGCTGGCTGTGACCGAGTGC 471
QY      126 ArgGluMetLeuHisGlnIleIleGlnAArgHisLeuThrAlaLysSerHisGlyArgVal 145
DB      472 CGGAGATGTGCTGTAGAGTGTGGAACACCACTTCAGGCCCAAGTACATGCGCGCATC 531
QY      146 AsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAlaLeuTyraAsnPrope 165
DB      532 CGCAGATGTGTTGATCATCTTCTGTGACCGAGTGTGCTACGCGCCCTCTATGGGCT--- 588
QY      166 GlyAsnPheLysProHisLeuGlnIlyLeuCysAspGlyIleAsnLysMetLeuAspGlu 185
DB      589 ---GACTTCACCTACGACCTTGGCAAGATCTGTGACGAGCTCAGAACTGCTAAGACGAA 645

RESULT 23
US-09-796-692-3207
/ Sequence 3207, Application US/09796692
/ Publication No. US20020198362A1
/ GENERAL INFORMATION:
/ APPLICANT: Galger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Manion, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
/ FILE REFERENCE: 2077.001200
/ CURRENT APPLICATION NUMBER: US/09/796,692
/ PRIOR FILING DATE: 2001-03-01
/ PRIOR APPLICATION NUMBER: 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ NUMBER OF SEQ ID NOS: 9597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-796-692-3207

Alignment Scores:
Pred. No.: 2,01e-47 Length: 287
Score: 504.00 Matches: 95

```

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 52.39% Indels: 0
 DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x US-09-796-692-3207 (1-287)

QY 74 ||leutyrtarqanbnanglnphasnqlnaspqluenualeuemetgluylsphenyls 93
 |||
 DB 3 ATTCTTAATGGAATTAATCAAGTTAATCAAGATGACGATTCATGATGAGAAATTTAAG 62
 QY 94 Lvelvvalhlsqlnuenualeuemetthrvalvalserphenlsqlnuvalaaptyrrthphe 113
 |||
 DB 63 AAGAAGTTCAATCAGCTTGCTATGACCGTGTCACTTCATCAGTGAGATTATACCTTT 122
 QY 114 Asparqanvalleuserarqleuenuaenqlucysargqlumetleuhsqlnlle 133
 |||
 DB 123 GACCGAATGTGTTATCCAGGCTGTTAAATGACAGAGATGCTGCACCAATCATTT 182
 QY 134 GlnarqhsleuthralaaysSerHlsqlyarqvalasnbnvalpheasphlsphenSer 153
 |||
 DB 183 CAGCGCACCTCACTGCACAGTCACATGACGCGGTTAATATGTTGATCATTTTCA 242
 QY 154 AspCyeglupheleu1a1aleutyrtanpProphcglyasnphe 168
 |||
 DB 243 GATTGTAAATTTTGGCTGCTGTATTAATCTTTTGGAAATTTT 287

RESULT 24

US-10-040-862-3207
 ; Sequence 3207, Application US/10040862
 ; Publication No. US20030078396A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Corixa Corporation
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
 ; FILE REFERENCE: 014058-013520US
 ; CURRENT APPLICATION NUMBER: US/10/040,862
 ; CURRENT FILING DATE: 2001-11-06
 ; PRIOR APPLICATION NUMBER: US 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: US 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: US 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: US 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: US 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: US 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; PRIOR APPLICATION NUMBER: US 09/796,692
 ; PRIOR FILING DATE: 2001-03-01
 ; NUMBER OF SEQ ID NOS: 10467
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3207
 ; LENGTH: 287
 ; TYPE: DNA

ORGANISM: Homo sapiens
 US-10-040-862-3207

Alignment Scores:

Score: 2,01e-47 Length: 287
 504.00 Matches: 95
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 52.39% Indels: 0
 DB: 14 Gaps: 0

US-10-627-571-2 (1-188) x US-10-040-862-3207 (1-287)

QY 74 ||leutyrtarqanbnanglnphasnqlnaspqluenualeuemetgluylsphenyls 93
 |||
 DB 3 ATTCTTAATGGAATTAATCAAGTTAATCAAGATGACGATTCATGATGAGAAATTTAAG 62
 QY 94 Lvelvvalhlsqlnuenualeuemetthrvalvalserphenlsqlnuvalaaptyrrthphe 113
 |||
 DB 63 AAGAAGTTCAATCAGCTTGCTATGACCGTGTCACTTCATCAGTGAGATTATACCTTT 122
 QY 114 Asparqanvalleuserarqleuenuaenqlucysargqlumetleuhsqlnlle 133
 |||
 DB 123 GACCGAATGTGTTATCCAGGCTGTTAAATGACAGAGATGCTGCACCAATCATTT 182
 QY 134 GlnarqhsleuthralaaysSerHlsqlyarqvalasnbnvalpheasphlsphenSer 153
 |||
 DB 183 CAGCGCACCTCACTGCACAGTCACATGACGCGGTTAATATGTTGATCATTTTCA 242
 QY 154 AspCyeglupheleu1a1aleutyrtanpProphcglyasnphe 168
 |||
 DB 243 GATTGTAAATTTTGGCTGCTGTATTAATCTTTTGGAAATTTT 287

RESULT 25

US-10-057-475B-3207
 ; Sequence 3207, Application US/10057475B
 ; Publication No. US2004002068A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Clapper, Jonathan David
 ; APPLICANT: Wang, Aljun
 ; APPLICANT: Ordenez, Nadia
 ; APPLICANT: Carter, Lauren
 ; APPLICANT: McNeill, Patricia Dianne
 ; APPLICANT: Corixa Corporation
 ; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
 ; FILE REFERENCE: 014058-014402US
 ; CURRENT APPLICATION NUMBER: US/10/057,475B
 ; CURRENT FILING DATE: 2002-01-22
 ; PRIOR APPLICATION NUMBER: US 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: US 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: US 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: US 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: US 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: US 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: US 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: US 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; Remaining Prior Application data removed - See File Wrapper or PALM.

```
/ NUMBER OF SEQ ID NOS: 10979
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-475B-3207
```

```
Alignment Scores:
Pred. No.: 2,01e-47 Length: 287
Score: 504.00 Matches: 95
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.39% Indels: 0
DB: 17 Gaps: 0
```

US-10-627-571-2 (1-188) x US-10-057-475B-3207 (1-287)

```
QY 11leuTYrArGAsnAnngInPheAnGlnAspGluLeuAlaLeuMetGluYspPheLys 93
Db 3 ATCTTATAGAAATATATCATGTTTATCAAGATGAGCTAGCATTTGATGGAAATTTAAG 62
QY 94 LysIysValHISGlnLeuAlaMetThrValValSerPheHISGlnValAspTYrThrPhe 113
Db 63 AAGAAAGTTCATCAGCTTCGTATGACCGGTGTCAGTTTCATCAGGTGATTAACCTTT 122
QY 114 AspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHISGlnIleIle 133
Db 123 GACCGAAATGTGTTATCCAGGCTGTTAATGAAATGACAGATGCTGCCAATAATCATTT 182
QY 134 GlnArgHISLeuThrAlaLysSerHISGlnYArgValAsnAsnValPheAspHISpHeSer 153
Db 183 CAGCGCACCTCATCTGCAAGTCACATGACGCGGTTAATATGTTGATCATTTTTC 242
QY 154 AspCysGluPheLeuAlaAlaLeuTYrAsnProPheGluYAsnPhe 168
Db 243 GATGTGAATTTTGGCTGCTGTATATCCTTTGGGAATTTT 287
```

RESULT 26

```
US-10-154-884B-3207
/ Sequence 3207, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
```

```
/ Remaining Prior Application data removed - See File Wrapper or PAM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-154-884B-3207
```

```
Alignment Scores:
Pred. No.: 2,01e-47 Length: 287
Score: 504.00 Matches: 95
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.39% Indels: 0
DB: 17 Gaps: 0
```

US-10-627-571-2 (1-188) x US-10-154-884B-3207 (1-287)

```
QY 74 11leuTYrArGAsnAnngInPheAnGlnAspGluLeuAlaLeuMetGluYspPheLys 93
Db 3 ATCTTATAGAAATATATCATGTTTATCAAGATGAGCTAGCATTTGATGGAAATTTAAG 62
QY 94 LysIysValHISGlnLeuAlaMetThrValValSerPheHISGlnValAspTYrThrPhe 113
Db 63 AAGAAAGTTCATCAGCTTCGTATGACCGGTGTCAGTTTCATCAGGTGATTAACCTTT 122
QY 114 AspArgAsnValLeuSerArgLeuLeuAsnGluCysArgGluMetLeuHISGlnIleIle 133
Db 123 GACCGAAATGTGTTATCCAGGCTGTTAATGAAATGACAGATGCTGCCAATAATCATTT 182
QY 134 GlnArgHISLeuThrAlaLysSerHISGlnYArgValAsnAsnValPheAspHISpHeSer 153
Db 183 CAGCGCACCTCATCTGCAAGTCACATGACGCGGTTAATATGTTGATCATTTTTC 242
QY 154 AspCysGluPheLeuAlaAlaLeuTYrAsnProPheGluYAsnPhe 168
Db 243 GATGTGAATTTTGGCTGCTGTATATCCTTTGGGAATTTT 287
```

RESULT 27

```
US-10-764-324-3207
/ Sequence 3207, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ CURRENT FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
```

```
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PAM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 3207
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-3207

Alignment Scores:
Pred. No.: 2,01e-47      Length: 287
Score: 504.00           Matches: 95
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.39%      Indels: 0
DB: 19                  Gaps: 0

US-10-627-571-2 (1-188) x US-10-764-324-3207 (1-287)

QY 74 ILEUTYARGAASNGINPHEASGLNAPGLULEUALALEUWETGIULYSPHELYS 93
DB 3 ATCTTTATGGAATATATCATGTTTAATCAAGATGAGCTTGAATGAGAAATTTAAG 62
QY 94 LVELYVALHISGLINLEUALAMETTRVALVASERPHISGLINVALASPTVYRTHRPH 113
DB 63 AAGAAAGTTCATCAGCTTGCTATGACCGTGTCAGTTTCATCAGTGGAATTAACCTTT 122
QY 114 ASPAASVALLEUSERARGLEULEASNGIUCYASRGILUMETLEUHSGLINLEILE 133
DB 123 GACCGGAATGTGTATTCACGCGTTTAAATGACAGAGATGCTGCACCAATCATTT 182
QY 134 GLNARGHISLEUTHRALALYSERHSISGLYARGVALASNAVALPHEASPHISPHESER 153
DB 183 CACCGCCACCTCACTGCCAAGTCACATGACGCGTTAATATGTGTTGATCATTTTCA 242
QY 154 ASPCYEGLUPHEULALALALEUTYRAANPROPHGLYASNPHE 168
DB 243 GATTGGAATTTTGGCTGCTGTATTAATCTTTGGGAATTTT 287

RESULT 28
US-09-796-692-2801/C
; Sequence 2801, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
; FILE REFERENCE: 2077, 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
```

```
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-2801

Alignment Scores:
Pred. No.: 7.37e-47      Length: 287
Score: 499.00           Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87%      Indels: 0
DB: 9                  Gaps: 0

US-10-627-571-2 (1-188) x US-09-796-692-2801 (1-287)

QY 74 ILEUTYARGAASNGINPHEASGLNAPGLULEUALALEUWETGIULYSPHELYS 93
DB 285 ATCTTTATGGAATATATCATGTTTAATCAAGATGAGCTTGAATGAGAAATTTAAG 226
QY 94 LVELYVALHISGLINLEUALAMETTRVALVASERPHISGLINVALASPTVYRTHRPH 113
DB 225 AAGAAAGTTCATCAGCTTGCTATGACCGTGTCAGTTTCATCAGTGGAATTAACCTTT 166
QY 114 ASPAASVALLEUSERARGLEULEASNGIUCYASRGILUMETLEUHSGLINLEILE 133
DB 165 GACCGGAATGTGTATTCACGCGTTTAAATGACAGAGATGCTGCACCAATCATTT 106
QY 134 GLNARGHISLEUTHRALALYSERHSISGLYARGVALASNAVALPHEASPHISPHESER 153
DB 105 CACCGCCACCTCACTGCCAAGTCACATGACGCGTTAATATGTGTTGATCATTTTCA 46
QY 154 ASPCYEGLUPHEULALALALEUTYRAANPROPHGLYASNPHE 168
DB 45 GATTGGAATTTTGGCTGCTGTATTAATCTTTGGGAATTTT 1

RESULT 29
US-10-040-862-2801/C
; Sequence 2801, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
```



```
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR FILING DATE: 2001-03-01
/ PRIOR FILING DATE: 2001-03-01
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2801
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-040-862-2801

Alignment Scores:
Pred. No.: 7,37e-47 Length: 287
Score: 499.00 Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87% Indels: 0
DB: 14 Gaps: 0

US-10-627-571-2 (1-188) x US-10-040-862-2801 (1-287)

QY 74 l l e l e u t t y r a r g a s n a n g l i n p h e a n g l i n a s p g l u l e u l a l e u m e t g l u l y s p h e l y s 93
DB 285 a t t c t t t a a g g a a t a t c a g t t t a a t c a a g a t g a c a t t g a t g a g a a a t t t a a g 226
QY 94 l y s l y s v a l i h i s g l i n e u l a m e t h r v a l s e r p h e i s g l i n v a l a s p y t r t h r p h e 113
DB 225 a a g a a a g t t c a t c a c t t c t t a t a c c g c t g t t a a t g a a t g a g a n t g c t g c a c c a a t c a t t 166
QY 114 a s p a r g a s n a l l e u s e r a r g l e u l e u a n g l u c y a r g l u m e t l e u h i s g l i n l e i l e 133
DB 165 g a c c g g a a t g t t a t t c c a g c c t g t t a a t g a a t g a g a n t g c t g c a c c a a t c a t t 106
QY 134 g l n a r g h i s l e u t h r a l a l y s e r h i s g l y a r g v a l a s n a s n v a l p h e a s p h i s p h e s e r 153
DB 105 c a g c g c c a c c t c a c t c c a a g t c a c a t g a c a g g g t t a a t a t g t t t g a t c a t t t t t t c a 46
QY 154 a s p c y s g l u p h e l e u l a l a l e u t y r a s n p r o p h e g l y a s n p h e 168
DB 45 g a t t g t g a a t t t t t t t g c t g c c t t g t a t a t c t t t t t g g a a t t t t 1

RESULT 30
US-10-057-475B-2801/c
/ Sequence 2801, Application US/10057475B
/ Publication No. US20040002068A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Clapper, Jonathan David
/ APPLICANT: Wang, Aijun
/ APPLICANT: Ordonez, Nadia
/ APPLICANT: Carter, Lauren
/ APPLICANT: McNeill, Patricia Dianne
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-01440205
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
```

```
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10979
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2801
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-475B-2801

Alignment Scores:
Pred. No.: 7,37e-47 Length: 287
Score: 499.00 Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-057-475B-2801 (1-287)

QY 74 l l e l e u t t y r a r g a s n a n g l i n p h e a n g l i n a s p g l u l e u l a l e u m e t g l u l y s p h e l y s 93
DB 285 a t t c t t t a a g g a a t a t c a g t t t a a t c a a g a t g a c a t t g a t g a g a a a t t t a a g 226
QY 94 l y s l y s v a l i h i s g l i n e u l a m e t h r v a l s e r p h e i s g l i n v a l a s p y t r t h r p h e 113
DB 225 a a g a a a g t t c a t c a c t t c t t a t a c c g c t g t t a a t g a a t g a g a n t g c t g c a c c a a t c a t t 166
QY 114 a s p a r g a s n a l l e u s e r a r g l e u l e u a n g l u c y a r g l u m e t l e u h i s g l i n l e i l e 133
DB 165 g a c c g g a a t g t t a t t c c a g c c t g t t a a t g a a t g a g a n t g c t g c a c c a a t c a t t 106
QY 134 g l n a r g h i s l e u t h r a l a l y s e r h i s g l y a r g v a l a s n a s n v a l p h e a s p h i s p h e s e r 153
DB 105 c a g c g c c a c c t c a c t c c a a g t c a c a t g a c a g g g t t a a t a t g t t t g a t c a t t t t t t c a 46
QY 154 a s p c y s g l u p h e l e u l a l a l e u t y r a s n p r o p h e g l y a s n p h e 168
DB 45 g a t t g t g a a t t t t t t t g c t g c c t t g t a t a t c t t t t t g g a a t t t t 1

RESULT 31
US-10-154-884B-2801/c
/ Sequence 2801, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-01352105
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
```

```

; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-154-884B-2801

Alignment Scores:
Pred. No.: 7.37e-47 Length: 287
Score: 499.00 Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-154-884B-2801 (1-287)

QY 74 ILEUTYRARGANBANGINPHEASNGIASPGLULEUALALEUMETGIULYSPHELYS 93
DB 285 ATCTTTATAGGGATATACGTTTATATCATGATGAGCTAGCATGATGAGAAATTTAAG 226
QY 94 LYLALYVAIHISGLINLEUALAMETTRVALYASERPHENISGLINVALASPYRTTHPHE 113
DB 225 AAGAAATTCATACGCTGTGATATGACCGGTGCTAGTTCCATGACGTGGATTAACCTTT 166
QY 114 ASPARGANVALLEUSERARGLEULEASNGIUCYASRGUWETLEUHSIGLILIE 133
DB 165 GACCGGATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATCATTT 106
QY 134 GLAARGHISLEUTHRALYSESRHISGLYARGLVALASMANVALPHEASPHISPHESER 153
DB 105 CAGCGCACCTCATCTGCACAGTCACATGACGCGGTATATATGTTGATCATTTTCA 46
QY 154 ASPCYAGLUPHELEUALAALALEUTYRASNPROPHG1YASNPHE 168
DB 45 GATTGTGAATTTTGGCTGCTGTATATATCCTTTTGGGAATTTT 1

RESULT 32
US-10-764-324-2801/c
; Sequence 2801, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10764,324
; PRIOR FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
```

```

; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-764-324-2801

Alignment Scores:
Pred. No.: 7.37e-47 Length: 287
Score: 499.00 Matches: 94
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.95% Mismatches: 0
Query Match: 51.87% Indels: 0
DB: 19 Gaps: 0

US-10-627-571-2 (1-188) x US-10-764-324-2801 (1-287)

QY 74 ILEUTYRARGANBANGINPHEASNGIASPGLULEUALALEUMETGIULYSPHELYS 93
DB 285 ATCTTTATAGGGATATACGTTTATATCATGATGAGCTAGCATGATGAGAAATTTAAG 226
QY 94 LYLALYVAIHISGLINLEUALAMETTRVALYASERPHENISGLINVALASPYRTTHPHE 113
DB 225 AAGAAATTCATACGCTGTGATATGACCGGTGCTAGTTCCATGACGTGGATTAACCTTT 166
QY 114 ASPARGANVALLEUSERARGLEULEASNGIUCYASRGUWETLEUHSIGLILIE 133
DB 165 GACCGGATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATCATTT 106
QY 134 GLAARGHISLEUTHRALYSESRHISGLYARGLVALASMANVALPHEASPHISPHESER 153
DB 105 CAGCGCACCTCATCTGCACAGTCACATGACGCGGTATATATGTTGATCATTTTCA 46
QY 154 ASPCYAGLUPHELEUALAALALEUTYRASNPROPHG1YASNPHE 168
DB 45 GATTGTGAATTTTGGCTGCTGTATATATCCTTTTGGGAATTTT 1

RESULT 33
US-09-796-692-5724
; Sequence 5724, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
; FILE REFERENCE: 2077.001200
; CURRENT APPLICATION NUMBER: US/09796,692
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
```

```
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ NUMBER OF SEQ ID NOS: 9597
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5724
/ LENGTH: 288
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (58)
/ OTHER INFORMATION: n=A,T,C or G
/ NAME/KEY: unsure
/ LOCATION: (59)
/ OTHER INFORMATION: n=A,T,C or G
US-09-796-692-5724

Alignment Scores:
Pred. No.: 3,65e-45      Length: 288
Score: 484.00          Matches: 94
Percent Similarity: 97.92%      Conservative: 0
Best Local Similarity: 97.92%      Mismatches: 1
Query Match: 50.31%      Indels: 1
DB: 9                  Gaps: 0

US-10-627-571-2 (1-188) x US-09-796-692-5724 (1-288)
QY 74 11leuTyTArGaBnAnGInPheAnGInAspGluLeuAlaLeuMetGluLys-PheLy 93
Db 3 ATTCTTATAGGAATATATCAAGTTTAATCAAGATGAGCTAGCATTTGATGGAAATNNCAA 62
QY 93 slyslYsValIHsGInLeuAlaMetThrValIserPheHsGInValAspTYrThrPh 113
Db 63 GAAGAAAGTTCACTCAGCTTCTATGACCGTGTCAAGTTTCCATCAGGTGATTTATACCTT 122
QY 113 eaBpArGaBnValleuSerArgLeuLeuAnGluCyAsrGluMetLeuHsGInIleI 133
Db 123 TGACCGGAATGTCTTATCCAGCGCTTTAAATGATGACAGATGCTGCACCAATCAT 182
QY 133 eGInArGhIsleuThAlaIalysSerHsGlyArgValAsnAsnValPheAspHisPhe 153
Db 183 TCAGCGCCACCTCACTGCCAAGTCACATGACGGGTTAATATATGTTGATCATTTTTC 242
QY 153 rAsPcYsGluPheLeuAlaIalauTYrAsnProPheGlyAsnPhe 168
Db 243 AGATTGGAATTTTGGCTGCTTGTATATCCTTTGGGAATTTT 288

RESULT 34
US-10-040-862-5724
/ Sequence 5724, Application US/10040862
/ Publication No. US20030078396A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
```

```
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: US 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR APPLICATION NUMBER: US 09/796,692
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5724
/ LENGTH: 288
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (58)
/ OTHER INFORMATION: n=A,T,C or G
/ NAME/KEY: unsure
/ LOCATION: (59)
/ OTHER INFORMATION: n=A,T,C or G
US-10-040-862-5724

Alignment Scores:
Pred. No.: 3,65e-45      Length: 288
Score: 484.00          Matches: 94
Percent Similarity: 97.92%      Conservative: 0
Best Local Similarity: 97.92%      Mismatches: 1
Query Match: 50.31%      Indels: 1
DB: 14                  Gaps: 0

US-10-627-571-2 (1-188) x US-10-040-862-5724 (1-288)
QY 74 11leuTyTArGaBnAnGInPheAnGInAspGluLeuAlaLeuMetGluLys-PheLy 93
Db 3 ATTCTTATAGGAATATATCAAGTTTAATCAAGATGAGCTAGCATTTGATGGAAATNNCAA 62
QY 93 slyslYsValIHsGInLeuAlaMetThrValIserPheHsGInValAspTYrThrPh 113
Db 63 GAAGAAAGTTCACTCAGCTTCTATGACCGTGTCAAGTTTCCATCAGGTGATTTATACCTT 122
QY 113 eaBpArGaBnValleuSerArgLeuLeuAnGluCyAsrGluMetLeuHsGInIleI 133
Db 123 TGACCGGAATGTCTTATCCAGCGCTTAAATGATGACAGATGCTGCACCAATCAT 182
QY 133 eGInArGhIsleuThAlaIalysSerHsGlyArgValAsnAsnValPheAspHisPhe 153
Db 183 TCAGCGCCACCTCACTGCCAAGTCACATGACGGGTTAATATATGTTGATCATTTTTC 242
```

QY	153	rAspCgSluPheLeuAlaIaleuTyrrAsnProPheGlyAsnPhe	168
Db	243	AGATTGGAAATTTTGGCTGCTGTATAATCCTTTTGGGAATTTT	288

RESULT 35

```

US-10-057-475B-5724
Sequence 5724, Application US/10057475B
Publication No. US2004002068A1
GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Aijun
APPLICANT: Ordonez, Nadia
APPLICANT: Carter, Lauren
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
of Infection: Hematological Malignancies
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5724
LENGTH: 288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)...(288)
OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-5724

```

Alignment Scores:	
Pred. No.:	3.65e-45
Score:	484.00
Percent Similarity:	97.92%
Best Local Similarity:	97.92%
Query Match:	50.31%
GB:	17
Length:	288
Matches:	94
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-10-627-571-2 (1-188) x US-10-057-475B-5724 (1-288)

QY 1 l l e n e t y r r y g s a n s n g i n p h e a n g i n a s p g i u l e u a l a l e m e g i l y s - P h e l y 93

Db 3 A T T C T T A T T G G A T A T A C A G T T A A T C A A G T G A C T G C A T T G A T G A T G A G A A T T N C A A 62

QY 93 s l y s V a l H i s G l n e u a l m e c h r V a l s e r P h e H i s g i n V a l s p y r t h n p h 113

Db	63	GAAGAAAGTTCACTACAGCTTGCTATACCGCTGGCTCACTTCATCAGGCGATTTATACCTT	1223
Qy	113	EAAPAPAGANVal1IeuSerArgLeuLeuAsnIuCyArgIuMetLeuHISGInIleI	133
Db	123	TGACCGGATGCTTTATCCAGCCTTTAAATGAATGCAGAGATGCTCGCCACCAATCAT	189
Qy	133	egIaARGHISLeuThrAlaIysSerHisGlyArgAlaAsnAsnValPheAspHisPheSe	155
Db	183	TCAGCGCACCTACTACGCCAAGTCAATGAGAGGGGCTTAAATATGCTTTGATCACTTTTC	244

Db 183 TCAGCGCACCTCACTGCCAAGTCACATGGA CGGTTAATAATGTGTTGATCATTTTC 242

QY

```

Db          243 AGATTGTGAATTTTGGCTGCTGTGTATTAATCCCTTTGGGAATTTT 288

RESULT 36
; Sequence 5724, Application US/10154884B
; Publication No. US2004005561A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154, 884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5724

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1...128)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-5724

```

US-10-627-571-2 (1-188) X US-10-154-884B-5724 (1-288)

QY 74 ILeuTYrARgAsnAsnGlnPheAsnGlnAspGluLeuAlaLeuMetGluLys-PheLY 93

```

Db      3  ATCTTTATAGAGATATACAGTATTATCAAGATGACATGATGAGAAATNNCA 62
Qy      93  slysyysvalhlsghlnleuallamethrvalvalserphenlsghlnvalaspyrthrph 113
Db      63  GAAGAAAGTTCACTACCTGCTATGACCGGTGTCACGTTTCCATCAGATGATATACCTT 122
Qy      113  eAspArGaenValleuSerArgleuLeuengluCyAsrgluMetleuHlsghlnlel 133
Db      123  TCACCGAATGTATTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATCAT 182
Qy      133  eglNArghlsleuThralalysSerHlsghlyArGvalAsnAsnValPheAspHisPhe 153
Db      183  TCAGCGCCACCTCAGTCCAGTCAATGACATGACGGGTTAAATGATGTTGATCATTTTTC 242
Qy      153  rAspCyGsluPheleuAlaaleuTyrrAsnProphleglyAsnpe 168
Db      243  AGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 288

```

RESULT 37

```

US-10-764-324-5724
/ Sequence 5724, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mamion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ PRIOR FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 5724
/ LENGTH: 288

```

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: unsure

/ LOCATION: (58)

/ OTHER INFORMATION: n=A,T,C or G

/ FEATURE:

/ NAME/KEY: unsure

/ LOCATION: (59)

/ OTHER INFORMATION: n=A,T,C or G

US-10-764-324-5724

Alignment Scores: 3.65e-45 Length: 288

```

Score: 484.00 Matches: 94
Percent Similarity: 97.92% Conservative: 0
Best Local Similarity: 97.92% Mismatches: 1
Query Match: 50.31% Indels: 1
DB: 19 Gaps: 0

```

US-10-627-571-2 (1-188) x US-10-764-324-5724 (1-288)

```

Qy      74  lleuTyrrArGaenValleuSerArgleuLeuengluCyAsrgluMetleuHlsghlnlel 93
Db      3  ATCTTTATAGAGATATACAGTATTATCAAGATGACATGATGAGAAATNNCA 62
Qy      93  slysyysvalhlsghlnleuallamethrvalvalserphenlsghlnvalaspyrthrph 113
Db      63  GAAGAAAGTTCACTACCTGCTATGACCGGTGTCACGTTTCCATCAGATGATATACCTT 122
Qy      113  eAspArGaenValleuSerArgleuLeuengluCyAsrgluMetleuHlsghlnlel 133
Db      123  TCACCGAATGTATTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATCAT 182
Qy      133  eglNArghlsleuThralalysSerHlsghlyArGvalAsnAsnValPheAspHisPhe 153
Db      183  TCAGCGCCACCTCAGTCCAGTCAATGACATGACGGGTTAAATGATGTTGATCATTTTTC 242
Qy      153  rAspCyGsluPheleuAlaaleuTyrrAsnProphleglyAsnpe 168
Db      243  AGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 288

```

RESULT 38

```

US-09-918-995-35521
/ Sequence 35521, Application US/09918995
/ Publication No. US20030073623A1
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
/ FILE REFERENCE: 20411-756
/ CURRENT APPLICATION NUMBER: US/09/918,995
/ PRIOR FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: US/09/235,076
/ PRIOR FILING DATE: 1999-01-20
/ NUMBER OF SEQ ID NOS: 38054
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 35521
/ LENGTH: 422
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-918-995-35521

```

Alignment Scores:

```

Pred. No.: 7.34e-27 Length: 422
Score: 324.00 Matches: 62
Percent Similarity: 78.95% Conservative: 28
Best Local Similarity: 54.39% Mismatches: 24
Query Match: 33.68% Indels: 0
DB: 10 Gaps: 0

```

US-10-627-571-2 (1-188) x US-09-918-995-35521 (1-422)

```

Qy      6  PheAsnSerlyAsnleuAlaValGlnAlaGlnlylslylleuGlylysketValser 25
Db      77  TTCAGCTCAAGAGAGCTGCACTGCACAGAGAAAGAACTACTAGTAAGATGCGGGT 136
Qy      26  lysSerllealArthrThreulleAspAspHrSerSerGluValleuAspGluLeuTyrr 45
Db      137  CGCTCTGTGGCTCATCTCTTCATATGATGACACACAGTGAAGTCTCTAGTACCTTAC 196
Qy      46  ArgValThrArGsluTyrrThrGlnAsnlyslsGlnAlaGlnlylslylleuAsnleu 65
Db      197  CGTGTCTCAGAGATACCGCACAGCGGCGCCAGCGCCGCTGATCAAGACCTG 256
Qy      66  lleuTyrrVallelleuLeuAlaalleuTyrrArGaenValleuPheAsnGlnAspGlu 85

```

Db 257 ATCAAGTGGCCATCAAGTGTGTGTGTCACCGCAATGCTCTTGTGCCAGTGAG 316
Qy 86 LeuAlaIeuMetGluYrPheIleValHisGlnIleuAlaMetThrValIleSer 105
Db 317 CTGACCTGTGCTACCGCTTTCGACAGAGCTGGGAGGCTGCCACGACCTTACG 376
Qy 106 PheHisGlnValaAspGlyrThrPheAspAlaGlnValIleuser 119
Db 377 TTGGTGTAGGTAGACTTTCACCTTCGAGGCTGCTGTCTGCT 418
RESULT 39
US-10-029-386-9211
; Sequence 9211, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9211
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC035144.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: NT HIT: g14725421, EVALUATE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BG18783.1, EVALUATE 0.00e+00
US-10-029-386-9211
Alignment Scores:
Pred. No.: 1,71e-24 Length: 500
Score: 304.00 Matches: 55
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 31.60% Indels: 0
Gaps: 0
DB: 16
US-10-627-571-2 (1-188) x US-10-029-386-9211 (1-500)
Qy 134 GlnArgHisLeuThrAlaIleSerHisGlyrValIleAsnValPheAspHisPheSer 153
Db 3 CAGGCGCACCTCATGCGCAAGTGCATGAGCGGTTATATGTGTGATCATTTTCA 62
Qy 154 AspCysGluPheLeuAlaAlaLeuTyraAspProPheGlyAsnPheIleProHisLeuGln 173
Db 63 GATTGTAATTTTGGCTGCTGCTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAA 122
Qy 174 LysIleuCysAspGlyIleAsnIleMetLeuAspGlnIleuSer 188
Db 123 AAACATATGTATGTATCAACAAATGTGATGAAGAACAATA 167
RESULT 40
US-10-972-079-232
; Sequence 232, Application US/10972079
; Publication No. US20050153317A1
; GENERAL INFORMATION:
; APPLICANT: MMT GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: ROSENFELD, David
; APPLICANT: KERR, Richard
; APPLICANT: BATES, Stephen
; APPLICANT: HOLM, Tom

; TITLE OF INVENTION: METHODS & SYSTEMS FOR INFERRING TRAITS TO BREED & MANAGE NON-BEE
; FILE REFERENCE: M01110-2
; CURRENT APPLICATION NUMBER: US/10/972,079
; CURRENT FILING DATE: 2004-10-22
; PRIOR APPLICATION NUMBER: US 60/514,333
; PRIOR FILING DATE: 2003-10-24
; NUMBER OF SEQ ID NOS: 96631
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 232
; LENGTH: 565
; TYPE: DNA
; ORGANISM: Chicken 19866894084041_1
US-10-972-079-232
Alignment Scores:
Pred. No.: 1.95e-11 Length: 565
Score: 189.00 Matches: 36
Percent Similarity: 88.46% Conservative: 10
Best Local Similarity: 69.23% Mismatches: 6
Query Match: 19.65% Indels: 0
Gaps: 0
DB: 22
US-10-627-571-2 (1-188) x US-10-972-079-232 (1-565)
Qy 4 AspValPheAsnSerIleAsnLeuAlaValGlnAlaGlnIleValIleLeuGlyIleMet 23
Db 409 GATTGTTTCAGTTCACAGATCTTGACATGCCAGCCAGAAAAAGATCTTGAATGAG 468
Qy 24 ValSerIleSerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeuAspGlu 43
Db 469 GCAACCAAAACCAATGCTATCATGCTCATGATGACACGACGACCAAAATCTTGAATGAG 528
Qy 44 LeuTyraGlyThrArgGlyrThrGlnAsnIle 55
Db 529 CTGTACAGGTGCACAAAGAACACACACCAAGAAACAA 564
RESULT 41
US-10-191-803-330/c
; Sequence 330, Application US/10191803
; Publication No. US20040014040A1
; GENERAL INFORMATION:
; APPLICANT: MENDRICK, Donna
; APPLICANT: PORTER, Mark
; APPLICANT: JOHNSON, Kory
; APPLICANT: HIGGS, Brandon
; APPLICANT: CASTLE, Arthur
; APPLICANT: ELASHOFF, Michael
; FILE REFERENCE: Cardiotoxin Molecular Toxicology Modeling
; CURRENT APPLICATION NUMBER: US/10/191,803
; CURRENT FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: US 60/303,819
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: US 60/305,623
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/369,351
; PRIOR FILING DATE: 2002-04-03
; PRIOR APPLICATION NUMBER: US 60/377,611
; PRIOR FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 1140
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 330
; LENGTH: 493
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040014040A1 AJ229843
US-10-191-803-330
Alignment Scores:
Pred. No.: 1.09e-07 Length: 493
Score: 155.00 Matches: 28

Percent Similarity: 96.55% Conservative: 0
Best Local Similarity: 96.55% Mismatches: 1
Query Match: 16.11% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-191-803-330 (1-493)

QY 160 AAlaLeuTyRAsnProPhelGlyAsnPhelYsProHisLeuGlnYsLeuCyAspGlyIle 179
Db 491 GCCTTGTATATATCCCTTTGGAAATTTAAACCTCAGAACTTTGTGACGGCATC 432

QY 180 AsnYsMetLeuAspGluGluAsnIle 188
Db 431 AACAAATGTGTGATGAAGAACATTA 405

RESULT 42

US-10-029-386-11708
Sequence 11708, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: A60MICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 11708
LENGTH: 543
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC012678.1
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.51
OTHER INFORMATION: EST HUMAN HIT: BG107777.1, EVALUE 0.00e+00
OTHER INFORMATION: NT HIT: AB002705.1, EVALUE 5.20e-01
OTHER INFORMATION: SWISSPROT HIT: P45815, EVALUE 3.90e+00

US-10-029-386-11708

Alignment Scores:

Pred. No.: 7,76e-07 Length: 543
Score: 148.00 Matches: 24
Percent Similarity: 88.64% Conservative: 15
Best Local Similarity: 54.55% Mismatches: 5
Query Match: 15.38% Indels: 0
DB: 16 Gaps: 0

US-10-627-571-2 (1-188) x US-10-029-386-11708 (1-543)

QY 143 GlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeuTyR 162
Db 3 GGGCGCATCAACACGCTTTACCACTTTCGAGATGTGAGTTCCTCCACCTCTAT 62

QY 163 AsnProPhelYsAsnPhelYsProHisLeuGlnYsLeuCyAspGlyIleAsnYsMet 182
Db 63 AGCTGTGATGAGACACTGTAGGCCCAACCTCAAGAGATTGTGAGGAATCAATTAAGTTG 122

QY 183 LeuAspGluGlu 186
Db 123 CTGATGAGAAA 134

RESULT 43

US-09-796-692-4917/C
Sequence 4917, Application US/09796692
Publication No. US20020198362A1
GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY

TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
FILE REFERENCE: 2077,001200
CURRENT APPLICATION NUMBER: US/09/796,692
CURRENT FILING DATE: 2001-03-01
PRIOR APPLICATION NUMBER: 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: 60/223,378
PRIOR FILING DATE: 2000-08-07
NUMBER OF SEQ ID NOS: 9597
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4917
LENGTH: 538
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (77)
OTHER INFORMATION: n=A,T,C or G

US-09-796-692-4917

Alignment Scores:

Pred. No.: 0.0192 Length: 538
Score: 109.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.33% Indels: 0
DB: 9 Gaps: 0

US-10-627-571-2 (1-188) x US-09-796-692-4917 (1-538)

QY 169 LysProHisLeuGlnYsLeuCyAspGlyIleAsnYsMetLeuAspGluGluAsnIle 188
Db 538 AAACCCCACTTCAAAAACATGATGATGATCAACAAATGTGTGAGAGAACATA 479

RESULT 44

US-10-040-862-4917/C
Sequence 4917, Application US/10040862
Publication No. US20030078396A1
GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17


```

; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (77)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-4917

Alignment Scores:
Pred. No.: 0.0192 Length: 538
Score: 109.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.33% Indels: 0
DB: 14 Gaps: 0

US-10-627-571-2 (1-188) x US-10-040-862-4917 (1-538)

QY 169 LysEProHisIleuGlnIysLeuCyAspGlyIleAsnIysMetLeuAspGluGluAsnIle 188
Db 538 AAACCCCACTTACAAAATGATGTGATGATCAACAAAATGTTGATGAAGAACATA 479

RESULT 45
US-10-057-475B-4917/C
; Sequence 4917, Application US/10057475B
; Publication No. US2004002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordonez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
```

```

; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(538)
; OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-4917

Alignment Scores:
Pred. No.: 0.0192 Length: 538
Score: 109.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.33% Indels: 0
DB: 17 Gaps: 0

US-10-627-571-2 (1-188) x US-10-057-475B-4917 (1-538)

QY 169 LysEProHisIleuGlnIysLeuCyAspGlyIleAsnIysMetLeuAspGluGluAsnIle 188
Db 538 AAACCCCACTTACAAAATGATGTGATGATCAACAAAATGTTGATGAAGAACATA 479

Search completed: July 28, 2005, 19:36:11
Job time : 671 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 28, 2005, 19:13:42 ; Search time 2996 Seconds

(without alignments)
2388.546 Million cell updates/sec

Title: US-10-627-571-2

Perfect score: 962

Sequence: 1 MATDVNSKRLAVQAQKTL.....KPHLOKLCDCGIMLDEBN1 188

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Egapop 6.0 , Egapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 100 summaries

Command line parameters:
-MODE=frame+p2n.model -DRV=x1h
-O=/cgm2.1/USPTO.spool.h/US10627571/runat.27072005.154720.27509/app.query.fasta_1.327
-DB=EST -OPMT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPEC=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=100
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10627571.@CGN.1.1.3437.@runat.27072005.154720.27509 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_g861: *
9: gb_g862: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	954	99.2	737	7	CN3053329	CN3053329 170006000
2	954	99.2	780	6	CB990165	CB990165 AGENCOURT
3	954	99.2	806	7	CF593784	CF593784 AGENCOURT
4	954	99.2	813	4	B1819214	B1819214 603034630
5	954	99.2	920	7	CF272384	CF272384 AGENCOURT
6	950	98.8	712	7	CN305327	CN305327 170005326
7	950	98.8	731	6	CD521721	CD521721 AGENCOURT
8	950	98.8	1055	5	BM919223	BM919223 AGENCOURT
9	949	98.6	718	1	AL549492	AL549492 AL549492

10	949	98.6	885	1	AL550457	AL550457
11	949	98.6	897	1	AL527566	AL527566
12	949	98.6	1610	3	CR611795	CR611795 full-length
13	949	98.6	1733	3	CR615621	CR615621 full-length
14	949	98.6	1754	3	CR615131	CR615131 full-length
15	947	98.4	896	5	BQ424670	BQ424670 AGENCOURT
16	947	98.4	998	7	CN648259	CN648259 ILLUMINAGEN
17	945	98.2	1045	1	AL554168	AL554168
18	934	97.1	779	7	CR548728	CR548728 DKEZP469E
19	933	97.0	876	4	BI752550	BI752550 603021969
20	927	96.4	773	5	BP433798	BP433798 BP433798
21	926	96.3	669	9	CN789950	CN789950 4124483 B
22	924	96.0	942	6	CA984413	CA984413 AGENCOURT
23	923	95.9	893	4	BI759027	BI759027 603042814
24	920	95.6	832	2	BP607236	BP607236 MY1_00017
25	920	95.6	841	6	CA469176	CA469176 AGENCOURT
26	920	95.6	865	4	BI525793	BI525793 602924756
27	918	95.4	559	1	AL558873	AL558873
28	916	95.4	580	5	BP302623	BP302623 BP302623
29	916	95.2	651	4	BQ146795	BQ146795 mab95H09.
30	916	95.2	816	4	AU135377	AU135377 AU135377
31	914	95.0	1542	3	AK090316	AK090316 Mus muscu
32	912	94.8	908	4	BI688217	BI688217 603314718
33	909	94.5	766	4	BI683693	BI683693 603306241
34	907	94.3	772	4	BI103517	BI103517 602889058
35	907	94.3	785	6	CB991862	CB991862 AGENCOURT
36	905	94.1	1836	3	AK087478	AK087478 Mus muscu
37	904	94.0	622	6	CD701806	CD701806 EST18330
38	902	93.8	656	4	BI687787	BI687787 603314107
39	901	93.7	707	7	CR557522	CR557522 DKEZP469E
40	900	93.6	566	7	CV023086	CV023086 5463 Full
41	898	93.3	779	1	AL542926	AL542926
42	893	92.8	747	6	CD520304	CD520304 AGENCOURT
43	892	92.7	671	6	BY743355	BY743355
44	891	92.6	739	2	BF140519	BF140519 601787536
45	889	92.4	692	7	CR771358	CR771358 DKEZP469E
46	885	92.0	711	4	BG972935	BG972935 602840605
47	885	92.0	834	5	BK405197	BK405197 BK405197
48	884	91.9	580	5	BP228279	BP228279
49	884	91.9	580	3	AK090350	AK090350 Mus muscu
50	879.5	91.4	757	6	CD520319	CD520319 AGENCOURT
51	879	91.4	697	1	AV713586	AV713586 AV713586
52	874	90.9	596	6	CA843219	CA843219 i555C03.Y
53	874	90.9	673	5	BP459271	BP459271 BP459271
54	872.5	90.7	593	7	CV027255	CV027255 5464 Full
55	868	90.2	762	7	CN305328	CN305328 170005999
56	865	89.9	762	5	BU124386	BU124386 603148153
57	863	89.7	588	6	CD723203	CD723203 0119007.Y
58	863	89.7	837	4	BI690706	BI690706 603314314
59	856	89.0	867	4	BG534704	BG534704 602553867
60	852	88.6	548	1	AA116718	AA116718 mg24d02.r
61	850	88.4	648	4	BG146696	BG146696 mab94h09.
62	848	88.1	863	4	BI686259	BI686259 603312767
63	844	87.7	572	7	CK903031	CK903031 i555C03.X
64	844	87.7	600	6	CA526496	CA526496 8012-74 M
65	842.5	87.6	942	7	CK805254	CK805254 AGENCOURT
66	840.5	87.4	676	1	AL642789	AL642789 AL642789
67	840.5	87.4	835	1	BX718081	BX718081 BX718081
68	840.5	87.4	864	5	BX733894	BX733894 BX733894
69	840.5	87.4	900	5	BX743252	BX743252 BX743252
70	840	87.3	784	4	CG640021	CG640021 OST372269
71	839.5	87.3	849	5	BX849289	BX849289 BX849289
72	838	86.9	615	6	CR771295	CR771295 DKEZP469M
73	836	86.1	605	6	CB554575	CB554575 NMSPO071
74	835	85.8	739	4	CF737286	CF737286 UT-M-HD0-
75	825.5	85.7	788	7	CR444882	CR444882 602506586
76	824.5	85.7	788	7	CR444882	CR444882 602506586
77	818.5	85.1	688	6	BY743855	BY743855
78	816	84.8	744	4	BI760235	BI760235 603045171
79	814	84.6	671	6	CB551156	CB551156 NMSPO066
80	812.5	84.5	640	6	CD639093	CD639093 AGENCOURT
81	809	84.1	607	6	CB552439	CB552439 NMSPO048
82	808	84.0	863	6	CD522233	CD522233 AGENCOURT

83	807	83.9	622	6	CB553311	MMSP0007
84	801	83.3	588	7	CR787140	DKPZ04695
85	799.5	83.1	771	6	CD521069	AGENCOURT
86	799.5	83.1	829	7	CK804285	AGENCOURT
87	799	83.1	674	2	BB667314	BB667314
88	795	82.6	821	1	AJ451661	AJ451661
89	795	82.6	975	2	BF781663	BF781663
90	794	82.5	652	6	CD638280	AGENCOURT
91	793	82.4	583	5	BP345938	BP345938
92	793	82.4	648	7	CN053290	Salamande
93	789	82.0	681	6	CD683714	EST234 hu
94	782	81.3	909	2	BE913315	60166522
95	780	81.1	661	7	CF737920	UI-M-HD-
96	778	80.9	766	1	AJ453373	AJ453373
97	777.5	80.8	648	1	AL888174	AL888174
98	772	80.2	582	5	BP261675	BP261675
99	770.5	80.1	680	4	BJ068640	BJ068640
100	757	78.7	517	6	CD372026	UI-R-GOO-

ALIGNMENTS

RESULT 1
CN305329 737 bp mRNA linear EST 16-MAY-2004
LOCUS 1700600027114 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN305329
ACCESSION CN305329
VERSION CN305329.1 GI:47321743

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 737)

Brandenberger, R., Wei, H., Zhang, S., Lei, S., Mursge, J., Fisk, G. J.,

Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M. S., Mandalam, R.,

Lebkowski, J. and Stanton, L. W.

Transcriptome characterization elucidates signaling networks that

control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6), 707-716 (2004)

Contact: Brandenberger R

Regenerative Medicine

230 Constitution Drive, Menlo Park, CA 94025, USA

Tel: 650 473 8658

Fax: 650 473 7760

Email: rbrandenberger@regeron.com

Insert Length: 737 Std Error: 0.00.

Location/Qualifiers

1..737

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/issue_type="embryonic stem cells, DMSO-treated H9 cell

line"

/note="Toigo dT primed, full-length enriched cDNA library

from DMSO-treated hES cell line H9 (p22) maintained in

feeder-free conditions"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-627-571-2 (1-188) x CN305329 (1-737)

QY

1

MetalThAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20

|||||

1

2

Db 107 ATGGCCACGANGTCTTAATTCGAAACAACTGGCGTTACAGCA CAAAGAGATCTTG 166

QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40

Db 167 GGTAAATGTGTCCAAATCCATCCGACACACCTTAATAGAGACAAAGTAGTAGGG 226

QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLysLysGluAlaGluLys 60

Db 227 CTGGATGACCTTACAGAGTGCACGAGAGTACACCCAAACAAAGAGGCAAGANG 286

QY LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyArgAsnAngin 80

Db 287 ATCATCAAGAACCTCATCAAGACAGATGCATCAAGCTGGCCATCTTTATAGAAATATCG 346

QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100

Db 347 TTTAATCAAGATGACCTGACATGATGAGAAATTAAGAAAGAAATTCATCAGCTTGT 406

QY 101 MetThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120

Db 407 ATGACCGTGTGTCAGTTTCCATCAGGTGATTTACCTTTGACCGGAATGTGTTATCCAG 466

QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140

Db 467 CTGTTAAATGAATGCAGAGATGCTGCACCAATCATTCACGCGCACCTCAGTCCAG 526

QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160

Db 527 TCACATGACGGGTAAATGATGTTGATCTTTTCAGATTTGAAATTTTGGCTGCG 586

QY 161 LeuTyAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180

Db 587 TTGTATTAATCCTTTTGGGAATTTTAAACCCACTTACAAAACCTATGATGATGATCAAC 646

QY 181 LysMetLeuAspGluGluAsnIle 188

Db 647 AAAATGTTGATGAAGAGAAACATA 670

RESULT 2

CN3090165 780 bp mRNA linear EST 01-MAY-2003

LOCUS AGENCOURT 13902271 NIH MGC 147 Homo sapiens cDNA clone

DEFINITION IMAGE:30347855 5', mRNA sequence.

ACCESSION CN3090165

VERSION CN3090165.1 GI:30284685

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 780)

NIH-MGC http://mgs.mci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

JOURNAL

COMMENT

Email: cgaabs-remail.nih.gov

Tissue Procurement: Dr. Stefan Hansson

cDNA library preparation: Michael J. Brownstein (NHGRI) with help

and advice from Piero Carninci (RIKEN)

cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/ILNL at:

http://image.llnl.gov

Plate: NDAM389 row: j column: 24

High quality sequence stop: 635.

Location/Qualifiers

1..780

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30347855"

/tissue_type="Human Placenta"

FEATURES

source

US-10-627-571-2 (1-188) x CN305329 (1-737)

QY

1

MetalThAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeu 20

|||||

1

2

/lab host="DH10B Tona"
 /clone lib="NIH_MGC_147"
 /note="Organ: placenta; Vector: pBluescriptR; Site 1:
 all-XhoI; Site 2: BamH; Oligo-dt primed using primer
 5'-TTTTTTTTTTTTT-3', size-selected for average
 insert size 2.3 kb and normalized to ROP 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: This is
 a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.:	7.26e-105	Length:	780
Score:	954.00	Matches:	187
Percent Similarity:	99.47%	Conservative:	0
Best Local Similarity:	99.47%	Mismatches:	1
Query Match:	99.17%	Indels:	0
DB:	6	Gaps:	0

US-10-627-571-2 (1-188) x CF593784 (1-780)

```

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLleu 20
DB 137 ATGCCCAAGATGCTTTAATTCAGAAACCTGGCCGTTTCAGGACAAAGAGATCTTG 196
QY 21 GlySerMetValSerLysSerLysLeaAlaThrThrLeuLeaAspThrSerSerGluVal 40
DB 197 GGTAAATGGTGTCCAATCATCATGCCACCCCTTATATAGCAGACAGATGAGAGTG 256
QY 41 LeuAspGluLeuTyraGValThrArgGluTyrrThrGlnAsnLysLysGlnAlaGlnLys 60
DB 257 CTGGATAGCTCTACAGAGTACAGGAGGATACCCCAAAACAGAGAGGAGGAGAG 316
QY 61 LysLleLysAsnLeuLysLysThrValLleLysLeuAlaLleLysTyraGAsnAsnGln 80
DB 317 ATCATCAAGAACCTCATCAAGAGATGCTACAGGCTGCTTATATAGGAATATACAG 376
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 377 TTTATATCAAGATGAGCTGATGATGATGAGAAATTTAAGAAAGATTCATGCTGCT 436
QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLleSerArg 120
DB 437 ATGACCTGTGTCATGCTTCATCATGATGATTAATACCTTTGACCGAATGTTATCCAG 496
QY 121 LeuLeuAsnGlnLysCysArgGluMetLeuHisGlnLleLleGlnArgHisLysThrAlaLys 140
DB 497 CTGTTAAATGATGACAGAGATGCTACCAAAATCATTCAGCCGCTCATCTGCGCAG 556
QY 141 SerHisGlnLysArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 557 TCACATGAGAGGGGTATATATGCTTGCATCATTTTTCAGATTCGTAATTTTGGCTGCC 616
QY 161 LeuTyraAsnProPheGlnAsnLysPheLysPheHisLysGlnLysLysCysAspGlyLleAsn 180
DB 617 TTGATATATCTCTTTTGGAAATTTTAAACCCACTTACAAAAACATATGATGATCAAC 676
QY 181 LysMetLeuAspGlnGluAsnLle 188
DB 677 AAAATGTTGATGAGAGAACATTA 700

```

RESULT 3
 CF593784 806 bp mRNA linear EST 26-SEP-2003
 LOCUS ACENECURT15624053 NIH_MGC_147 Homo sapiens cDNA clone
 DEFINITION IMAGE:30528465 5', mRNA sequence.
 ACCESSION CF593784
 VERSION CF593784.1 GI:36347693
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 1. (bases 1 to 806)
 AUTHORS
 NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 Unpublished (1999)
 COMMENT
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgabs-rc@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hansson
 cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: NDAM614 row: h column: 10
 High quality sequence stop: 648.
 Location/Qualifiers

FEATURES

source

1. 806
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30528465"
 /tissue_type="Human Placenta"
 /lab host="DH10B Tona"
 /clone lib="NIH_MGC_147"
 /note="Organ: placenta; Vector: pBluescriptR; Site 1:
 all-XhoI; Site 2: BamH; Oligo-dt primed using primer
 5'-TTTTTTTTTTTTT-3', size-selected for average
 insert size 2.3 kb and normalized to ROP 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Carninci, in
 preparation). Library constructed by M. Brownstein
 (NIH/NHGRI, National Institutes of Health). Note: This is
 a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.:	7.59e-105	Length:	806
Score:	954.00	Matches:	187
Percent Similarity:	99.47%	Conservative:	0
Best Local Similarity:	99.47%	Mismatches:	1
Query Match:	99.17%	Indels:	0
DB:	7	Gaps:	0

US-10-627-571-2 (1-188) x CF593784 (1-806)

```

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLysLleu 20
DB 138 ATGCCCAAGATGCTTTAATTCAGAAACCTGGCCGTTTCAGGACAAAGAGATCTTG 197
QY 21 GlySerMetValSerLysSerLysLeaAlaThrThrLeuLeaAspThrSerSerGluVal 40
DB 198 GGTAAATGGTGTCCAATCATCATGCCACCCCTTATATAGCAGACAGATGAGAGTG 257
QY 41 LeuAspGluLeuTyraGValThrArgGluTyrrThrGlnAsnLysLysGlnAlaGlnLys 60
DB 258 CTGGATAGCTCTACAGAGTACAGGAGGATACCCCAAAACAGAGAGGAGGAGAG 317
QY 61 LysLleLysAsnLeuLysLysThrValLleLysLeuAlaLleLysTyraGAsnAsnGln 80
DB 318 ATCATCAAGAACCTCATCAAGAGATGCTACAGGCTGCTTATATAGGAATATACAG 377
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 378 TTTAATCAAGATGAGCTGATGATGAGAAATTTAAGAGAAATTCATCATGCTTGGCT 437
QY 101 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValLleSerArg 120

```

Db 438 ATGACCGGTGTCAGTTTCATCATCAGTGGATTATACCTTGGACCGGAGTGTATTCAGG 497
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 498 CTGTTAAATGATATCGAGAGATGCTGCACCAATATTCAGGCCACTTCAGTCCAG 557
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
 Db 558 TCACATGAGACGGGTATATATATGCTTTCATCATTTTTCAGATTGTGATTTTGGCTGCC 617
 QY 161 LeuTyrAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 618 TTGTATTAATCCTTTTGGGAAATTTTAAACCCACTTCAAAACTATGTGATGTATCAAC 677
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 678 AAAATGTTGGATGAGAGAACATA 701
 RESULT 4
 B1819214 813 bp mRNA linear EST 04-OCT-2001
 LOCUS 603034630F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175702 5',
 DEFINITION mRNA sequence.
 ACCESSION B1819214
 VERSION B1819214.1 GI:15930764
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 813)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 COMMENT
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM1437 row: 1 column: 07
 High quality sequence stop: 809.
 Location/Qualifiers
 1..813
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5175702"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 Kb,
 insert size range 1-3 Kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

Pired. No.: 7.68e-105 Length: 813
 Score: 954.00 Matches: 187
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 1
 Query Match: 99.17% Indels: 0
 Db: 4 Gaps: 0

US-10-627-571-2 (1-188) x B1819214 (1-813)
 QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysLeu 20
 Db 98 ATGGCCACGAGATGCTTTTAAATTCAAACCTGGCCGTTCAAGCCAAAGAAATCTTG 157
 QY 21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 158 GGTAAATAGTGTCCAAATCCATCCGCCACCTTAATGAGACAAAGTGTAGG 217
 QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLeuGlnAlaGluLys 60
 Db 218 CTGATGAGCTTACAGAGTGCACGAGGATGCACCCAAACAAAGAGGACAGAAAG 277
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgSerAsnGln 80
 Db 278 ATCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATATGGAATATACG 337
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 338 TTTAATCAAGATGACTGACATTGATGAGAAATTTAAGAAAGATTCATCAGCTTGT 397
 QY 101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 398 ATGACCGTGTGATGTTCCATCAGGTGATTTATACCTTGGACCGGAATGTATTCCAGG 457
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 458 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCCACTTCCTCCAG 517
 QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIle 160
 Db 518 TCACATGAGACGGGTATATATGCTTTCATCATTTTTCAGATTGTGAATTTTGGCTGCC 577
 QY 161 LeuTyrAsnProPheGlyAsnPhelysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 578 TTGTATTAATCCTTTTGGGAAATTTTAAACCCACTTCAAAACTATGTGATGTATCAAC 637
 QY 181 LysMetLeuAspGluGluAsnIle 188
 Db 638 AAAATGTTGGATGAGAGAACATA 661
 RESULT 5
 CF272384 920 bp mRNA linear EST 13-AUG-2003
 LOCUS CF272384
 DEFINITION AGENCOURT_15178772 NIH_MGC_192 Homo sapiens cDNA clone
 IMAGE:30512832 5', mRNA sequence.
 ACCESSION CF272384
 VERSION CF272384.1 GI:33628296
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 920)
 NIH-MGC http://mgs.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 COMMENT
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10K07 Bethesda, MD 20892
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Agencourt
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM573 row: m column: 01
 High quality sequence stop: 598.
 Location/Qualifiers

FEATURES

Source

1. .920
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30512832"
/tissue_type="Brain glioblastoma"
/lab_host="D110B (T1 phase-resistant)"
/clone_id="N1H.MGC.197"
/notes="Vector: pExpress1; Site 1: SmaI; Site 2: NotI; The library was constructed by reverse transcription of 1 ug mRNA using the oligo dt primer GCGGCGCCCTT20 and an RNaseH⁺ and MMLV reverse transcriptase. Second strand synthesis was carried out by standard methods. The cDNA was size selected by agarose gel for > 1.2 kb, digested with Not I and directionally cloned into the vector Express-1 at the SmaI/NotI sites. DNA from the primary library was used for in vitro transcription from the T7 promoter to produce biotinylated RNA transcripts. These biotinylated transcripts, along with blocking oligos to the poly-A, multiple cloning site and primer regions, were hybridized with single stranded circles produced by phage-mediated production from the primary library to a Cot value of 10-20. Streptavidin/phenol extraction was utilized to remove DNA:RNA hybrids leaving un-hybridized single stranded circles which were repaied by primer extension and transformed back into E. coli resulting in the normalized library. Average insert size 2.0 kb. 3' linker/adaptor sequence GCGGCGCCCTT20. This library was constructed by Agencourt Bioscience."

ORIGIN

Alignment Scores:	
Pred. No.:	9.05e-105
Score:	954.00
Percent Similarity:	99.47%
Best Local Similarity:	99.47%
Query Match:	99.17%
BB:	7
Length:	920
Matches:	187
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-10-627-571-2 (1-188) X CF272384 (1-920)

Qy		1	MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGluLysLeu1	leu	20
Dd		98	ATGGCCACAGATCTCTTTAATTCCAAAACCCTGCCCTTCAGGCACAAGAATCTTG		15
Qy		21	GlyLysMetValSerIysSer11ealathrIleu11eaSpBthrSerSeri1val		40
Dd		158	GGTAAATGGTGTCCNAATCCCATCGCCACCACTTAATAGACGACACAAGTAGTAGGTG		21
Qy		41	LeuAspGluLeuThrArgValThzArgGluYrrThrGlnAsnLysGluAlaGluLys		60
Dd		218	CTGAATGAGCTCTACAGAGTGAACAGGGGAGTACACCCAAAACMAAGAGGACAGAG		277
Qy		61	Lys11LeysAsnLeu11elysThrVal11elysLeuAla11eleutyArgAsnAsnGln		80
Dd		278	ATCATCAAGACTCATTCAAGACAGTCATCAAGCTGGCCATTTCTTAATAGAAATATCAC		333
Qy		81	PheAsnGlnAspGluLeuAlaleuMetGluLysPheLysLysValHisGlnLeuAla		100
Dd		338	TTTAAATCAAGATGAGCTTACATTGATGGGAATTTAAGAAAGATTCATCACCTTGCT		397
Qy		101	MetThrValValSerPheHisGlnValAspYrrThrPheAspArgAsnValLeuSerArg		120
Dd		398	ATGACCGTGGTCAGTTTCATCATGAGGAGTTATATACCTTTCACCGGAATGTGTATCCAG		457
Qy		121	IeuLeuAsnGluCySaArgGluMetLeuHisGlnIle11eglinaHis11leuThrAlaLys		140
Dd		458	CTGTAAATGAATGACGAGAGATGTCGACCAAAATCATTCAGCGCCACTCATCGCAAG		517
Qy		141	SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCyGluPheLeuAlaLa		160
Dd		518	TTCATGAGGACGGGTATATATAGCTTTGTATCATTTTTACAGTTGTGAATTTTTGGCTGCC		577

Qy	Db
161 LeuTYrAspMPcPhGhIYAspNheIyProhIleuGlnIyIleAsnCYsAspGlyIleAsn 160	578 TGTGTAATACCTTTGGGAAATTTAAACCCACTTACAAAACTATGTGATGTATCAAC 633
181 IyMeIleuAspGluGluAsnIle 188	
638 AAAATGTTGATGAAGAAGACATA 661	

RESULT 6

LOCUS	CN305327	712 bp	mRNA	linear	EST 16-MAY-2006
DEFINITION	CN305327	I7000532623591 GRN ES Homo sapiens	CDNA 5'	mRNA sequence.	
ACCESSION	CN305327				
VERSION	CN305327.1	GI:47321741			
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				

FEATURES

Source

ORIGIN

Alignment Scores:	
Pred. No.:	1,36e-104
Score:	950.00
Percent Similarity:	99.47%
Best Local Similarity:	98.94%
Query Match:	98.75%
BB:	7
Length:	712
Matches:	186
Conservative:	1
Mismatches:	1
Indels:	0
Gaps:	0

US-10-627-571-2 (1-188) X CN305327 (1-712)

Qy	1	metatatrAspVal.PheAnsSerIysAsnLeuAlaGlnIagIlybVslIleu	20
	:::		
Db	81	GGGGCCACAGATGTCTTTAATTCCAAACCTGGCCGTTCAGGCACAAAAGATCTTG	140
Qy	21	GIlySmecValSerIysSerIleAlaThrThrLeuIleAspThrSerSerGIuVal	40
Db	141	GGTAAATAGTGTCTCAATCATGTGCCACACCTTAATAGCAGCACAAATAGAGGTG	200
Qy	41	IeuAspGIuLeuTYrArgValIThrArgGIuTYrThrGlnAsnIySylGluIagIlybS	60
Db	201	CTGATGTAGCTCTACAGAGTACACACGGAGATACACCAAAACAAAGAGGACAGAGAAG	260
Qy	61	IySIIleYbAsnIleuIIeYbThrValIIeYbLeuAlaIleuTYrArgAsnAsnGln	80
Db	261	ATCATCAAGAACTTCATATAGACAGCATCAAGCTGGCCATCTCTTAATAGAAATAATCAG	320


```

QY      81 PheAenGlnAaspGluLeuAlaLeuMetGluLeuPheLysLysValHisGlnLeuAla 100
      |||
Db      321 TTTAAATCAAGATAGCTAGCATTCATGAGAAATTTAAAGAAAGTTATCATCGCTTGGT 380
QY      101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAnValIleuSerArg 120
      |||
Db      381 ATGACCGGTGGTCACTTCCATCCAGTGGATTATACCTTTGACCGGAATGTTATCCAGG 440
QY      121 LeuLeuAenGluCyAArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||
Db      441 CTTGTAAATGAAGCAGAGAGATGCTGCACCAATCATTCAGGCCACCTCAGTCCCAAG 500
QY      141 SerHisGlyArgValAsnAenValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
      |||
Db      501 TCACATGAGCGGCTTAATATATGTTGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 560
QY      161 LeuTyrAsnProDheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
      |||
Db      561 TTGTATTAATCTTTTGGGAATTTTAAACCCCACTTCAAAAACCTATGTGATGCTATCAAC 620
QY      181 LysMetLeuAaspGluGluAsnIle 188
      |||
Db      621 AAAATGTTGATGAAGAACATATA 644

RESULT 7
LOCUS   CD521721 731 bp mRNA linear EST 06-JUN-2003
DEFINITION AGENCOURT_1435034 NIH_MGC_191 Homo sapiens cDNA clone
IMAGE:30413002 5', mRNA sequence.
ACCESSION CD521721
VERSION   CD521721.1 GI:31453439
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL   Contact: Daniela S. Gerhard, Ph.D.
COMMENT   Office of Cancer Genomics
          National Cancer Institute / NIH
          Bldg. 31 Rm10A07 Bethesda, MD 20892
          Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: Narayan Bhat
          cDNA Library Preparation: CLONTECH Laboratories, Inc.
          DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
          DNA Sequencing by: Agencourt Bioscience Corporation
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LNL at:
          http://image.lnl.gov
          Plate: NDCM207 row: e column: 11
          High quality sequence stop: 603.
          Location/Qualifiers
            1..731
              /organism="Homo sapiens"
              /mol_type="mRNA"
              /db_xref="taxon:9606"
              /clone="IMAGE:30413002"
              /issue_type="Pooled"
              /lab_host="DH10B (TI phage-resistant)"
              /note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcaggcc);
              Site 2: SfiI (ggcgcctcgcc); Library is oligo-dT primed
              and directionally cloned. PBMC - Peripheral Blood
              Mononuclear Cells. RNA was pooled from 3/6hour stimulation
              with PMA adn Ionomycin. 5' and 3' adaptors were used in
              cloning as follows: 5' adaptor sequence:
              5'-CACGGCCATTATGCGC-3' and 3' adaptor sequence:
              5'-ATTCTAAGGCCGAGCGCGCGACATG-dt(30)BN-3' (where B = A,
              C, or G and N = A, C, G, or T). Average insert size 1.69
              kb (range 0.70-5.0 kb). 15/15 colonies contained inserts

```

ORIGIN

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC library."

Alignment Scores:

Pred. No.:	2,03e-104	Length:	731
Score:	950.00	Matches:	186
Percent Similarity:	99.47%	Conservative:	1
Best Local Similarity:	98.94%	Mismatches:	1
Query Match:	98.75%	Indels:	0
DB:	6	Gaps:	0

US-10-627-571-2 (1-188) x CD521721 (1-731)

```

QY      1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlGlnLysIleLeu 20
      |||
Db      30 GTGGCCACAGATGTCTTTAATTCAAAAACCTGGCCGTTCAAGCACAAGAGATCTTG 89
QY      21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      |||
Db      90 GGTAAGATGTCATCAAAATCCATCCGACACCTTAATAGACGACAAATGATGAGG 149
QY      41 LeuAaspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGluIaGluLys 60
      |||
Db      150 CTGGAATGAGCTTACAGAGTGACAGGGAGTACACCCAAAACAAGAGAGGACAGAAAG 209
QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
      |||
Db      210 ATCATCAAGACCTCATCAAGACAGTCAACAGCTGCGCATCTTATGTGAATATACAG 269
QY      81 PheAenGlnAaspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAla 100
      |||
Db      270 TTTAAATCAAGATAGCTTACATTTGATGAGAAATTTAAAGAAAGTTATCATCGCTTGGT 329
QY      101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAnValIleuSerArg 120
      |||
Db      330 ATGACCGGTGGTCACTTCCATCCAGTGGATTATACCTTTGACCGGAATGTTATCCAGG 389
QY      121 LeuLeuAenGluCyAArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      |||
Db      390 CTTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCACCTCAGTCCCAAG 449
QY      141 SerHisGlyArgValAsnAenValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
      |||
Db      450 TCACATGAGCGGCTTAATATATGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCC 509
QY      161 LeuTyrAsnProDheGlyAsnPhelysProHisLeuGlnLysLeuCyAspGlyIleAsn 180
      |||
Db      510 TTGTATTAATCTTTTGGGAATTTTAAACCCCACTTCAAAAACCTATGTGATGCTATCAAC 569
QY      181 LysMetLeuAaspGluGluAsnIle 188
      |||
Db      570 AAAATGTTGATGAAGAACATATA 593

RESULT 8
LOCUS   BM919223 1055 bp mRNA linear EST 12-MAR-2002
DEFINITION AGENCOURT_6715690 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748442
IMAGE:5748442
ACCESSION BM919223
VERSION   BM919223
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE     Unpublished (1999)
JOURNAL   Contact: Robert Strausberg, Ph.D.
COMMENT   Email: cgabbs-r@mail.nih.gov

```


QY 82 AaenGlnaSpGluLeuAlaLeuMeCgLuLySpheLyLyLyValHISglnLeuAlaMeC 101
 DB 243 AATCAAGATGAGCTACATTGATGAGGAATTTAAAGAGAAATTCATCGCTTGCCTATG 302
 QY 102 ThrValValSerPheHISglnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
 DB 303 ACCGTGTCAGTTTCATCGATCGATGATTAATACCTTTGACCGGAATGTGTATCCAGGCTG 362
 QY 122 LeuAsnGluCyArgGluMeCLeuHISglnLeuHISglnArgHISleuThrAlaLySer 141
 DB 363 TTAATGAATGACGAGAGATGCTGACCAATATCATTCAGCGCCACTCATGCGCAAGTCA 422
 QY 142 HISglYArgValAsnAsnValPheAspHISpHeSerAspCySglnPheLeuAlaAlaLeu 161
 DB 423 CATGACGGGTTAATATATGTTTGAATCATTTTCAGATGTGAAATTTTGGCTGCCTTG 482
 QY 162 TyrAsnProPheGluAsnPhelysProHISleuGlnLyLeuCyAspGlyIleAsnLyS 181
 DB 483 TATAAATCCTTTGGGAATTTTAAACCCCACTTCAAAAATATGTATGTATGATCAACAA 542
 QY 182 MetLeuAspGluGluAsnIle 188
 DB 543 ATGTTGATGAAGAGAACATA 563

RESULT 10
 AL550457 885 bp mRNA linear EST 25-MAR-2004
 LOCUS AL550457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION AL550457 5-PRIME, mRNA sequence.
 ACCESSION AL550457
 VERSION AL550457.3 GI:45750828
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 885)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31272274.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by life technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 3485.f

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CSODI057AA04QP1c=3485.f.
 FEATURES
 source
 1..885
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI057YA07"
 /issue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:

Pred. No.: 3,47e-104 Length: 885
 Score: 949.00 Matches: 186
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 1

Query Match: 98.65% Indels: 0
 DB: 1 Gaps: 0
 US-10-627-571-2 (1-188) x AL550457 (1-885)

QY 2 AlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLySISleuGly 21
 DB 3 GCCACGATGCTTTAATTCCAAAACCTGGCGTTACGGCACAAAAGAGATCTTGGGT 62
 QY 22 LyMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
 DB 63 AAAAAGGTGTCCAAATTCATCCACCACTTAAATAGACAGACAAAGATGATGAGTGTG 122
 QY 42 AaPGLuLeuTyrArgValThrArgLyuTyrThrGlnAsnLyLyGlnAlaGlnLyS 61
 DB 123 CATGACCTTCAGAGCTGACAGGAGTACACCCAAACAGAGAGGCGAGAGATGC 182
 QY 62 ILeYsAsnLeuIleYsThrValIleYsLeuAlaIleLeuTyrArgAsnGlnPhe 81
 DB 183 ATCAAGAACTCATCAAGACATTAATCAAGCTGGCCATCTTTATAGGATATACGTTT 242
 QY 82 AaenGlnaSpGluLeuAlaLeuMeCgLuLySpheLyLyLyValHISglnLeuAlaMeC 101
 DB 243 AATCAAGATGAGCTACATTGATGAGGAATTTAAAGAGAAATTCATCGCTTGCCTATG 302
 QY 102 ThrValValSerPheHISglnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
 DB 303 ACCGTGTCAGTTTCATCGATCGATGATTAATACCTTTGACCGGAATGTGTATCCAGGCTG 362
 QY 122 LeuAsnGluCyArgGluMeCLeuHISglnLeuHISglnArgHISleuThrAlaLySer 141
 DB 363 TTAATGAATGACGAGAGATGCTGACCAATATCATTCAGCGCCACTCATGCGCAAGTCA 422
 QY 142 HISglYArgValAsnAsnValPheAspHISpHeSerAspCySglnPheLeuAlaAlaLeu 161
 DB 423 CATGACGGGTTAATATATGTTTGAATCATTTTCAGATGTGAAATTTTGGCTGCCTTG 482
 QY 162 TyrAsnProPheGluAsnPhelysProHISleuGlnLyLeuCyAspGlyIleAsnLyS 181
 DB 483 TATAAATCCTTTGGGAATTTTAAACCCCACTTCAAAAATATGTATGTATGATCAACAA 542
 QY 182 MetLeuAspGluGluAsnIle 188
 DB 543 ATGTTGATGAAGAGAACATA 563

RESULT 11
 AL527566 897 bp mRNA linear EST 24-MAR-2004
 LOCUS AL527566 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION AL527566 cDNA clone CSODC024YA23 5-PRIME, mRNA sequence.
 ACCESSION AL527566
 VERSION AL527566.3 GI:45702664
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 897)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 13, 2001 this sequence version replaced gi:31065417.
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: sequef@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized. Library
 was constructed by life technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 3485.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?S=CS00C024AA120P1&c=3485.f.

FEATURES
Location/Qualifiers

1. .897

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS00C024YA23"

/issue_type="NEUROBLASTOMA COT 25-NORMALIZED"

/clone_id="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-627-571-2 (1-188) x AL527566 (1-897)

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .1610

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS00D1057YA07"

/issue_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT_6"

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-627-571-2 (1-188) x CR611795 (1-1610)

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

```

QY      162 TyrAsnProPheGlyAsnPhelysProHlsleuGlnLysLeuCyAspGlyIleAsnLys 181
      483 TATATCCCTTTGGGAATTTTAACCCCACTTACAAAACTAATGATGATATCAACAAA 542
DB
QY      182 MetLeuAspGluGlnAsnIle 188
      543 ATGTGGATGAAGAACATCA 563
DB
RESULT 13
CR615621 1733 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0DC024YA23 of Neuroblastoma Cot
DEFINITION 25-normalized of Homo sapiens (human).
ACCESSION CR615621
VERSION CR615621.1 GI:50496428
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1733)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1..1733
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC024YA23"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 8.5e-104 Length: 1733
Score: 949.00 Matches: 186
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 98.65% Indels: 0
DB: 3 Gaps: 0
US-10-627-571-2 (1-188) x CR615621 (1-1733)
QY      2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlaGlnLysIleLeuGly 21
      3 GCACAGATGCTCTTATATCCAAAACTGGCGCTTCAGGCACAAAAGAAAGATCTTGGGT 62
DB
QY      22 LysMetValSerLysSerLysIleAlaThrThrLeuIleAspAspThrSerSerGluValIleu 41
      63 AAAATGGTGTCCAAATCCATCGCACACCTTAATATGAGACCAAGTAGTGGGTGCTG 122
DB
QY      42 AspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGlnAlaGlnLysLys 61
      123 GATGAGCTCTTACAGAGTGCACAGGAGTGCACCCAAAACAGAGAGGACAGAGAGATC 182
DB
QY      62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLysTyrArgAsnGlnIle 81
      111 LeuLysAsnLeuIleLysThrValIleLysLeuAlaIleLysTyrArgAsnGlnIle 81

```

```

DB      183 ATCAAGAACTTCATCAAGACAGTCAATCAAGCTGCCATCTTTATAGGAATATCACTTT 242
QY      82 AsnGlnAspGluLeuAlaLeuMetGlnLysPheLysLysValHisGlnLeuAlaMet 101
      243 AATCAAGATGAGCTGACATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTCTTAG 302
DB
QY      102 ThrValIleSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArgLeu 121
      303 ACCGTGTCAGTTCATCAGTGCATTAACCTTTGACCGGAAATGTGTATCCAGGCTG 362
DB
QY      122 LeuAsnGluCyAspArgLysMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
      363 TTAATATGATGACAGAGAGATGCTGCACCAATCATTCAGCGCACCTCACTGCCAAGTCA 422
DB
QY      142 HisGluArgValAsnAsnValPheAspHisPheSerAspCyAspGlnPheLeuAlaLeu 161
      423 CATTGACGGGTTAATTAATGTTGTATCATCTTTTCAAGATGTGAATTTTGGCTCCCTTG 482
DB
QY      162 TyrAsnProPheGlyAsnPhelysProHlsleuGlnLysLeuCyAspGlyIleAsnLys 181
      483 TATATCCCTTTGGGAATTTTAACCCCACTTACAAAACTAATGATGATGATCAACAAA 542
DB
QY      182 MetLeuAspGluGlnAsnIle 188
      543 ATGTGGATGAAGAACATCA 563
DB
RESULT 14
CR615131 1754 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS0D1081Y011 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR615131
VERSION CR615131.1 GI:50495938
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1754)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1..1754
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1081Y011"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Alignment Scores:
Pred. No.: 8.64e-104 Length: 1754
Score: 949.00 Matches: 186
Percent Similarity: 99.47% Conservative: 0
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 98.65% Indels: 0
DB: 3 Gaps: 0

```

US-10-627-571-2 (1-188) x CR615131 (1-1754)

QY 2 AlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAgluLysIleLeuGly 21
 DB 3 GCCACAGATGCTTTAATCCAAAAACCTGGCCGTTGAGCACAAGAGATCTTGGT 62
 QY 22 LysMetValSerLysSerIleAlaThrThreuleuIleAspAspThrSerSerGluValLeu 41
 DB 63 AAATGGTGTCCAAATCATCGCACCACTTAATACACACACAGTAGAGAGTCTG 122
 QY 42 AspGluLeuValThrArgGluValThrGlnAsnLysGluValAgluLysLys 61
 DB 123 GATGAGCTTACAGAGGACGAGGAGTACACCCAAACAAAGAGGAGCAGACATC 182
 QY 62 IleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuValArgAsnGlnPhe 81
 DB 183 ATCAGAGACCTCATCAAGACAGTCATCAGCTGGCCATCTTATATAGAAATATCAGTTT 242
 QY 82 AsnGluAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAlaMet 101
 DB 243 AATCAAGATAGCTAGCATGATGAGAGAAATTAAGAGAAAGTTCAATCAGCTTGTATG 302
 QY 102 ThrValAlaSerPheHisGlnValAspValThrPheAspArgAsnValLeuSerArgLeu 121
 DB 303 ACCGTGATCAGTTTCCATCAGGTGATTAATACCTTTGACCGAATGTGTATCCAGGCTG 362
 QY 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
 DB 363 TTAAATCAATGCAGAGAGATGCTGCACCAATCATTTAGGCGCACCTCATCTGCAGATCA 422
 QY 142 HisGluValArgValAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
 DB 423 CAGGACGGGTATATATGTTGATTCATTTTCAGATTGTGATTTTGGCTGCTG 482
 QY 162 TyrAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleLeuLys 181
 DB 483 TATTAATCTTTGGGAATTTTAACCCCACTTACAAAACCTAATGTATGATCAACAA 542
 QY 182 MetLeuAspGluGluAsnIle 188
 DB 543 ATGTTGATGAGAGAGAACATTA 563
 RESULT 15
 BQ424670 896 bp mRNA linear EST 23-MAY-2002
 LOCUS AGENCOURT 7833747 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6153217
 DEFINITION 5', mRNA sequence.
 ACCESSION BQ424670
 VERSION BQ424670.1 GI:21119985
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 896)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bax-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM13492 row: f column: 02
 High quality sequence stop: 596.
 Location/Qualifiers
 1..896
 /organism="Homo sapiens"

/mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6153217"
 /issue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH MGC 67"
 /note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

ORIGIN

Alignment Scores:
 Pred. No.: 6.15e-104 Length: 896
 Score: 947.00 Matches: 186
 Percent Similarity: 98.94% Conservative: 0
 Best Local Similarity: 98.94% Mismatches: 2
 Query Match: 98.44% Indels: 0
 DB: 5 Gaps: 0

US-10-627-571-2 (1-188) x BQ424670 (1-896)

QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAgluLysIleLeu 20
 DB 97 ATGACACAGATGCTTTAATTCAAAACCTGGCCGTTTCAGGACCAAGAGATCTTG 156
 QY 21 GlyLysMetValSerLysSerIleAlaThrThreuleuIleAspAspThrSerSerGluVal 40
 DB 157 GGTAAATGGTGTCCAAATCATCGCACCACTTAATAGACACACAGTACTAGAGTG 216
 QY 41 LeuAspGluLeuValThrArgGluValThrGlnAsnLysGluValAgluLys 60
 DB 217 CTGGATGAGCTTACAGAGTACAGGAGTACACCAAAACAAAGAGGAGCAGAGAG 276
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuValArgAsnGln 80
 DB 277 ATCATCAAGACCTCATCAAGACATCATCAGCTGCCATCTTATAGAAATATCAG 336
 QY 81 PheAsnGluAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 DB 337 TTTAATCAAGATAGCTAGCATTTGATGAGAAATTTAAGAGAAATTCATCAGCTTGT 396
 QY 101 MetThrValAlaSerPheHisGlnValAspValThrPheAspArgAsnValLeuSerArg 120
 DB 397 ATGACCGGTGCTCATCTCATCAGTGTGATTAATCCTTTGACCGAATGTGTATCCAG 456
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 DB 457 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCACTGCTGCAG 516
 QY 141 SerHisGluValArgValAsnValPheAspHisPheSerAspCysGluPheLeuAlaLys 160
 DB 517 TCACATGACGCGGTATATATGTGTGATCATTTTTCAGATTGTAAATTTTGGCTGCC 576
 QY 161 LeuTyrAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 DB 577 TTGTAAATCTTTTNGAATTTTAAACCCCACTTCAAAACTAGTATGTATGATCAAC 636
 QY 181 LysMetLeuAspGluGluAsnIle 188
 DB 637 AAAATGTGATGAGAGAACATTA 660
 RESULT 16
 CN648259 998 bp mRNA linear EST 13-MAY-2004
 LOCUS ILLUMIGN_MGC_30108 Katze MRPB Macaca mulatta cDNA clone IBTMB:6795
 DEFINITION 5' similar to Bases 72 to 998 highly similar to human TNFAIP8
 (Hs.17839), mRNA sequence.
 ACCESSION CN648259
 VERSION CN648259.1 GI:47161702
 KEYWORDS EST.
 SOURCE Macaca mulatta (rhesus monkey)
 ORGANISM Macaca mulatta

REFERENCE
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Macaca.
1 (bases 1 to 998)
Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magnus, C.L.
Large-scale Rhesus Macaque cdna sequencing
Unpublished (2003)
Contact: C. Magnus

Eukaryota; Euteleostomi; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecidae; Cercopithecinae; Macaca.

1 (bases 1 to 998)

Katze, M.G., Thomas, M., Korth, M., Tadonato, S.P. and Magness, C.L.
Large-scale Rhesus Macaque cDNA Sequencing
Unpublished (2003)
Contact: C. Magness
illumigen Biosciences, Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2004.03.27. 798 Q20 bases.

PCR Primers
FORWARD: CCTCACTAAAGGGAACAATA
BACKWARD: CACTATAGGCGGATTTGGTA
Insert length: 998 Std Error: 0.00
Plate: CL000256 row: D column: 03
Seq primer: CCTCACTAAAGGGAACAATA
SOLTA=No.

FEATURES

Source

```
1. .998
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/clone="IBRW:6795"
/sex="male"
/tissue_type="blood"
/cell_type="PBMC"
/dev_stage="adult"
/lab_host="E. coli SOLR"
/clone_id="Katz MPB"
/notes="Vector: Uni-Zap XR; Site 1: EcoR I; Site 2: Xho I"
Created from Stragene zap-cDNA Synthesis kit Catalog
#200400 and Zap-CDNA Gigapack III Gold Cloning kit
Catalog #200450"
```

ORIGIN

Alignment Scores:	
Pred. No.:	7,11e-104
Score:	947.00
Percent Similarity:	99.47%
Best Local Similarity:	98.40%
Query Match:	98.44%
DB:	7
	Gaps: 0
	Indels: 0
	Mismatches: 1
	Conservative: 2
	Matches: 18
	length: 999

US-10-627-571-2 (1-188) X CN648259 (1-998)

QY	1	MetAlaThrAspVal.PheAsnSerLysAsnLeuAlaVal.GlnIaGlnLysLysIleLeu 20
	...	
Db	93	GTCGCACAGAGTCTTTATTTCCAAAAACCTGCCGTTCAAGCACAAGAAAGACTTG 152
QY	21	GlyLysMetVal.SerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
Db	153	GGTAAAAAGTGTCCTCAAAATCCATCCGCACACCTTATATGATACACGAGTATGAGGTC 212
QY	41	LeuAspGluLeuLysArgVal.ThrArgGluLysThrGlnAsnLysLysGlnIaGlnLys 60
Db	213	CTGAGTAGAGCTTACAGAGTGAACGAGGAGTACACCCAAACAGAAAGAGCGACAGAAC 272
QY	61	LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuLysArgAsnAsnGln 80
Db	273	ATCAATCAAGAACCTTATATCAAGACGGTCATCAAGCTGCGCATCTTTACAGGAATAATCAG 332
QY	81	PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysLysValHisGlnLeuAla 100
Db	333	TTTATATCAGAGAGACTGACCTGAGTGGAGAAATTTAAGAGAAAGTTTCATCAACTTGCT 392
QY	101	MetThrValVal.SerPheHisGlnValAspLysThrPheAspArgAsnValLeuSerArg 120
Db	393	ATGACCGGAGTACAGTTTCCACAGAGGAGATTTATACCTTTGACCGGAGATGTTATCCAGG 452

QY 122 LeuLeuAsnGluCyAArgGluMetLeuHisGlnLeilegGlnAArgHisLeuThrAlaLeu 140
Db 453 CTTTAAATGAATGCAAGAGATGCTGCACCAATTCATTCAAGCTCACTCATCTCCAG 512
QY 141 SerHisGlyArgValAsnAsnValPheAspHisSerSerAspCysGluPheLeuAla 160
Db 513 TCACATGACGAGGTAAATATATGCTTTGATCATTTTCAAGTTGTGACTTTTGGCTGCC 572
QY 161 LeuTyrAsnProPheGlyAsnPhelysProHisLeuGlnIlysLeuCysAspGlyIleAsn 180
Db 573 TTGTATAATCCCTTTGGGAAATTTTAAACCTCATTAACAAAACTATGTGATGATCAAC 632
QY 181 LysMetLeuAspGluGlnAsnIle 188
Db 633 AAAATGTTGATGAAGAGACATA 656

RESULT	17
AL554168	
LOCUS	AL554168 1045 bp mRNA linear EST 30-MAR-2004
DEFINITION	AL554168 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSOD1081Y011 5-PRIME, AL554168.
ACCESSION	AL554168
VERSION	AL554168.3 GI:4585929
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE 1 (bases 1 to 1045)
AUTHORS Li W.B., Grubler C., Jeesee J. and Polayev D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31275981

Contact: Genoscope
Genoscope - Centre National de Séquençage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five primer
ends enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalised. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3485.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?c=OS0DI081AHO60P1&c=3485.f>.

FEATURES

Source

```

/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:3606"
/clone="CS0D1081Y011"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Alignment Scores:	
Pred. No.:	1,32e-103
Score:	945.00
Percent Similarity:	98.93%
Best Local Similarity:	98.93%
Query Match:	98.23%
DB:	1
Gaps:	0
Length:	1045
Matches:	185
Conservative:	0
Mismatch:	2
Indels:	0
Gaps:	0

US-10-627-571-2 (1-188) X AL554168 (1-1045)

OY 2 AlaHisAspValPheasnSerLysASnIleuAlaValGlnAlaGlnLysLysIleLeucGly 21
Db 3 GCCACAGATGCTTTAAATTCCAAAACTGGCGCTTCAGGCACAAGAAGATCTTGGGT 62

QY	22	lysmeValSerlySer11eal1aThrhLeu1leapApThSeSeCglValleu	41
Db	63	AAAAAGGTGTCCAAATCCATGCCACCACTTAAATAGACACACAAAGTGTAGGTGTCTG	122
QY	42	AapGluLeuTyArgValThrArgGluTyThrGlnAenlySLyGluAglulLyLyS	61
Db	123	GATGAGCTCTACAGAGTACCGAGGAGTACACCCAAACCAAGAAAGGACAGAAAGATC	182
QY	62	IllyAsnleuIleLyThyValIleLySleuAalleuTyArgAsnAenGlnphe	81
Db	183	ATCAAGAACCTTCATCAAGACAGTMAATCAAGCTGGCATTTTAAAGAAATATCATCTT	242
QY	82	AenGlnAepGluLeuAleuMeGluLySPhelySLySLyValIleGluLeuA1aMet	101
Db	243	AATCAAGATGAGCTTGATGATGAGAAATTTAAGAAAGAAAGTTTCACTAGCTTGATG	302
QY	102	ThrValValSerPheIleGlnValAapTyThrPheAspArgAsnValleuSerArgleu	121
Db	303	ACCGGTCAGCTTTCATCAGGTGATTTATACCTTTGACCGGAATGTGTATCCAGCTG	362
QY	122	LeuAenGluCyArgGluMetLeuH1eGlnIleIleGlnA1aGh1SleuThralAlySer	141
Db	363	TTAAATGATGACAGAGATGCTGACCAATCATTTACGCCGCACTTCACTGCCAAGCA	422
QY	142	H1eGluArgValAAsnAsnValPheAspHisPheSerAapCyGluPheLeuA1aAlaLeu	161
Db	423	CATGACCGGTTTAAATATGTGTTTATCATTTTTCAGATTTGTGAATTTTGGCTGCTTG	482
QY	162	TyrAenProPheGluAsnPhelyBroHisleuGlnLySleuCyAspGlyIleAsnLyS	181
Db	483	TATATCTTTTGGGAATTTTAAACCCACCTTACAAAACATATGATGATGATCAACAA	542
QY	182	MetLeuAepGluGluAenIle	188
Db	543	ATGTTGATGAAGAGACATA	563
RESULT 18			
LOCUS	CR548728		
DEFINITION	DKFZp469B1232.r1.469 (synonym: pk1d1) Pongo pygmaeus cDNA clone		
ACCESSION	DKFZp469B1232.5, mRNA sequence.		
VERSION	CR548728		
KEYWORDS	CR548728.1 GI:50242352		
SOURCE	EST.		
ORGANISM	Pongo pygmaeus (orangutan)		
REFERENCE	Pongo pygmaeus		
AUTHORS	EubalYota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo. 1 (bases 1 to 779)		
TITLE	Bahr, A., Lauber, J., Mewes, H.W., Weill, B., Auld, C., Osanger, A., Foch, G., Han, M. and Wiemann, S.		
JOURNAL	Pongo pygmaeus mRNA (Bahr, A., Lauber, J., Mewes, H.W., et al.)		
COMMENT	Unpublished (2004)		
	Contact: MIPS		
	MIPS		
	Incoltaedter lander1.1, D-85764 Neuherberg, Germany		
	This is the 5' sequence of the clone insert clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Qigen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project. This clone (DKFZp469B1232) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de. Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/ .		
FEATURES	location/Qualifiers		
source	1..779		
	/organism="Pongo pygmaeus"		
	/mol_type="mRNA"		
	/db_xref="taxon:9600"		
	/clone="DKFZp469B1232"		

ORIGIN					/tissue_type="kidney" /dev_stage="adult" /lab_host="RDH10B" /clone_lib="469 (synonym: pklid)" /note="vector: pspori1_sfi; site_1: sfil1; site_2: sfilB"
Alignment Scores:					
Pred. No.:	1.92e-102	Length:	779		
Score:	934.00	Matches:	183		
Percent Similarity:	98.94%	Conservative:	3		
Best Local Similarity:	97.34%	Mismatches:	2		
Query Match:	97.09%	Indels:	0		
DB:	7	Gaps:	0		
US-10-627-571-2 (1-188) x CRS48728 (1-779)					
QY	1	MeCaLaThRaApVaLpheaSeriLyAsnLeuAlaValGlnAlaGlnLyLeIleu	20		
DB	184	GTGGCCACAGATGCTTTTAATTCGAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTG	243		
QY	21	GLySMeTValSerLySerLeAlaThThrLeuIleAspThrSerSerGluVal	40		
DB	244	GCTAAATAGTGTCCAATCATCATCGCCACCACTTAATAGACGACAGAGTAGAGGTG	303		
QY	41	LeuAspGluLeuTyraGValThraGluTyrrThrgInsnLySlySgluAlGluLyS	60		
DB	304	CTGGATGAGCTCTCAGAGTGACCGAGGAGATCACTCAAAACAAGAGAGCAGAGAG	363		
QY	61	LySleLyAsnLeuIleTyThValIleLyLeuAlaIleLeuTyraGAsnAsnGln	80		
DB	364	ATCATCAAGAACCTCATCAAGACAGTATCAAGCTGCCATCTTTATAGAAATATACG	423		
QY	81	PheaSngInAspGluLeuAlaLeuMeGluLySPhelySlySValHisGlnLeuAla	100		
DB	424	TTTATCAAGATGAGCTTNGCTTAATGAGAGAAATTTAAGAGAAAGTTCACTGCTTCT	483		
QY	101	MeThraValLySerPheHisGlnValAspTyrrThrPheaAspAsnValLeuSerArg	120		
DB	484	ATGACCGGTCGATTTCCATCAGGTGATTAATCCTTGACCGGAATGTGTATCCAGG	543		
QY	121	LeuLeuAsnGluCyBaArgGluMeLeuHisGlnIleIleGlnAlaGHisLeuThraLyS	140		
DB	544	CTGTTAATAGATGCAGAGAGATCTGCACCAATATCATTCGCGCTCACCTCACGCCAAG	603		
QY	141	SeriHisGlyArgValaAsnValPheAspHisPheSerAspCySgLuPheLeuAla	160		
DB	604	TCACATGACCGGTTAATATGCTTTGATCATTTTTCAGATGATGATTTTGGCTGCC	663		
QY	161	LeuTyraSnpPropheGlyAsnPhelySProHisLeuGlnLySLeuCySAspGlyIleasn	180		
DB	664	TTGATATATCTTTTGGGAATTTTAAACCCACATTACAAAACATATGATGATGATCAAC	723		
QY	181	LySMeLeuAspGluGluAsnIle	188		
DB	724	AAATGTGATGAAGAGACATA	747		
RESULT 19					
LOCUS	B1752550	876 bp	mRNA	linear	EST 25-SEP-2001
DEFINITION	603021966P1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5192561 5',				
ACCESSION	B1752550				
VERSION	B1752550.1				
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE	NIH-MGC http://mgc.nci.nih.gov/.				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-rxmail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LIML at:
 http://image.llnl.gov
 Plate: LHM11481 row: 1 column: 18
 High quality sequence stop: 813.
 Location/Qualifiers

FEATURES

source

```

1..876
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5192561"
/lab_host="NIH_MGC_114"
/clone_lib="NIH_MGC_114"
/notes="Organ: brain; Vector: PCMV-SPORT6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

```

ORIGIN

Alignment Scores:

Pred. No.:	2..966-102	Length:	876
Score:	933.00	Matches:	186
Percent Similarity:	98.41%	Conservative:	0
Best Local Similarity:	98.41%	Mismatches:	2
Query Match:	96.99%	Indels:	1
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x B1752550 (1-876)

```

QY 1 MetAlaThraSpValPheAsnSerLyAsnLeuAlaValGlnAglInlySylleu 20
DB 118 ATGGCCACAGATGTCCTTAATTCACAAAACCTGGCGGTCAGGACACAAAGATCTTG 177
QY 21 GlyLysMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 178 GGTAAATGCTGTCACAAATCCATGCCACCACTTATAGACACAAAGTAGTGAGTG 237
QY 41 LeuAspGluLeuTyraValThraGluTyThrGlnAsnLySylGlnAglInlySyl 60
DB 238 CTGGATGAGCTCTACAGAGTGACCAAGGAGTACACCAAAACAGAGAGGCGAGAG 297
QY 61 LysIleLyAsnLeuIleLyThrValIleLySleuAlaIleLeuTyraGlnAsnGln 80
DB 298 ATCATCAAAACCTCATCAAGACAGCATCAAGCTGGCATCTTATATAGATATATCAG 357
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLyPheLySylValIleGlnAla 100
DB 358 TTTAAACAAGATGAGCTTATGATGAGAAATTTAAGAGAAAGTTCAGCTTGCT 417
QY 101 MetThrValIleSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
DB 418 ATGACCGTGGTCAGTTTCATCATCGATGATTAACCTTTGACCGGATGTTATCCAG 477
QY 121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 478 CTGTTAAAGATGAGAGAGATGCTGCACCAATATTCAGGCGCACCTGCTGCAAG 537
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGlnPheLeuAla 160
DB 538 TACATCGACGGGTTATATATGTTGATCATTTTCAGATTTGTAATTTTGGGCTGCC 597
QY 161 LeuTyraAsnPro-PheGlyAsnPheLySylLeuGlnIleLyLeuCyAspGlyLeuAs 180

```

Db

598 TTGTAATATCCTTTTGGGATTTTAACCCCATTAACAAAACATATGATGATCA 657

QY

180 nlySwetLeuAspGluGlnAlaIle 188

Db

658 CAAAATGTTGATGAAGAGAACATA 682

RESULT 20

BP433798

BP433798 773 bp mRNA linear EST 30-DEC-2003

LOCUS

BP433798 full-length enriched swine cDNA library, adult lung Sus

DEFINITION

scrofa cDNA clone UMG010079A10 5', mRNA sequence.

ACCESSION

BP433798

VERSION

BP433798.1 GI:40423865

KEYWORDS

EST.

SOURCE

Sus scrofa (pig)

ORGANISM

Sus scrofa

REFERENCE

Uenishi, H., Bouchi, T., Suzuki, K., Sawazaki, T., Toki, D., Shinkai, H.,

AUTHORS

Okumura, N., Hamasima, N., and Awata, T.

TITLE

PEDF (Pig EST Data Explorer): construction of a database for ESTs

JOURNAL

Nucleic Acids Res. 32 (1), D484-D488 (2004)

COMMENT

Animal Genome Laboratory, Genome Research Department

Vector sequences were eliminated by RepeatMasker version 2002/07/13

Low quality bases were trimmed based on the quality values.

FEATURES

Location/Qualifiers

ORIGIN

Alignment Scores:

Pred. No.:	1..336-101	Length:	773
Score:	927.00	Matches:	179
Percent Similarity:	98.94%	Conservative:	7
Best Local Similarity:	95.21%	Mismatches:	2
Query Match:	96.36%	Indels:	0
DB:	5	Gaps:	0

US-10-627-571-2 (1-188) x BP433798 (1-773)

```

QY 1 MetAlaThraSpValPheAsnSerLyAsnLeuAlaValGlnAglInlySylleu 20
DB 154 GTGGCCACAGATGTCCTTAATTCACAAAACCTGGCGGTCAGGACACAAAGATCTTG 213
QY 21 GlyLysMetValSerLySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
DB 214 GGTAAAGTGTGTCACAAATCCATGCCACCACTTATAGACACAAAGATGATGCGAAG 273
QY 41 LeuAspGluLeuTyraValThraGluTyThrGlnAsnLySylGlnAglInlySyl 60
DB 274 CTGGATGAGCTCTACAGGAGTGACCAAGAGTACACCAAGAACAGAGAGGCGAGAAA 333

```

QY 61 Lysilelysaenleuilelythrvallilelyleuallilelytrargasaenlgn 80
 Db 334 ATCATCAAAACCTCATCAAAACGATCAAGCTGCCATTCTCAGAAATCAAA 393
 QY 81 Pheasnlnaspgluleuallaleuemetglulysphelyslyvalhsglnleualla 100
 Db 394 TTTTACCAGATGAGCTGACGATGATGAGAAATTTTAAAGAAAGTCCACCACTTGGCT 453
 QY 101 Metthrvalvalserpnehlsglnvalaspytrthrphesaparganvalleuserarg 120
 Db 454 ATGACCGGTGTCATGTTTTCATCAGGTGATTTTCCCTTTACCGGATGTGCTATCCAG 513
 QY 121 Leuleuansglucysargluleuemetleuhsiglnleileglnarghlsleuthrallys 140
 Db 514 CTGCTCATATGATGACAGAGCTGCTCCACAGATCATCCAGCTCATCCCGCCAG 573
 QY 141 SerHlsGlyArGyAlaenValPheasPhlsPheSerAspCySGlupheleuAla 160
 Db 574 TCACACGACGAGATTAACAAGCTTTGATCATTTTCAAGTTGTGATTTCTTGGCCGCC 633
 QY 161 LeuTyAsnProphelglsanphelysProHlsleuGlnlyseuCyAspGlylleasn 180
 Db 634 TTGTATATATCCCTTTGAAATTTCAACCCGACTTACAGAACTGTGTGATGTATCAAC 693
 QY 181 LysMetLeuAspGluGluAsnille 188
 Db 694 AAAATGTTGGATGAGATTAACATA 717

RESULT 21
 CN789950

LOCUS DEFINITION 4124483 BARC 8BOV Bos taurus cDNA clone 8BOV_35M20 5', mRNA
 sequence.
 CN789950

ACCESSION CN789950.1 GI:47685930

VERSION KEYWORDS

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

REFERENCE 1 (bases 1 to 669)
 Baumann,R.G., Baldwin,R.L., Sonstegard,T.S., Van Tassell,C.P. and Matukumalli,L.K.
 Construction and Analysis of a CDNA Library Generated From Intestinal Muscle and Epithelial Tissues of Holstein Cattle .Unpublished (2004)

JOURNAL Contact: Richard G. Baumann
 Bovine Functional Genomics Lab
 ANRI
 BUDG 162: BARC-EAST, Beltsville, MD 20705, USA
 Tel: 3015048604
 Fax: 3015048744
 Email: rbaumann@anri.barc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred 0.000925 using options -trim alt '' -trim fastavector identified by cross match using options -mismatch 12 -mismatch 12
 plate: 35 row: M column: 20
 Seq primer: CCTATTAGGTGACACTATAGAAC
 High quality sequence stop: 669.
 Location/Qualifiers
 1. 669
 /organism="Bos taurus"
 /mol_type="mRNA"
 /strain="Holstein"
 /db_xref="taxon:9913"
 /clone="8BOV_35M20"
 /sex="Female"
 /tissue_type="Epithelial, Muscle"
 /dev_stage="Lactating, Neonatal"
 /lab_host="DH10B TONK"
 /clone_id="BARC 8BOV"

Alignment Scores:
 Pred. No.: 1,45E-101 Length: 669
 Score: 926.00 Matches: 179
 Percent Similarity: 98.40% Conservative: 6
 Best Local Similarity: 95.21% Mismatches: 3
 Query Match: 96.26% Indels: 0
 DB: 7 Gaps: 0

US-10-627-571-2 (1-188) x CN789950 (1-669)

QY 1 MetAlthrAspValPheasnserysaenleuallilelytrargasaenlgn 20
 Db 90 GTGGCTACAGATGCTTTTAAATCCAAAACCTGCCGCTTCCAGACAAAAGATCTTG 149
 QY 21 GlyMetValSerlyserllealathrthrleuileaspherserarglval 40
 Db 150 GGCAAAATGGCATCCAGTCCATGCAACCCCTCATGACACAACAGCATGAGTG 209
 QY 41 LeuAspGluLeuTyArGyAlaThrArgGluTyThrGlnAsnlysglnlaGluTy 60
 Db 210 CTGATGAGCTCTACAGGCTGACCAAGATACACCAAGAAAGAGGCAAGAA 269
 QY 61 Lysilelysaenleuilelythrvallilelyleuallilelytrargasaenlgn 80
 Db 270 ATCATCAAGAACCTCATCAAGACCGTCATCAACCTGCCATCTTTAACAAGATTAAC 329
 QY 81 Pheasnlnaspgluleuallaleuemetglulysphelyslyvalhsglnleualla 100
 Db 330 TTTTACCAGACGAGCTGACATGATGAGAAATTTTAAAGAAAGTTCATCAGCTTGGCT 389
 QY 101 MetThrValValSerPhehlsglnvalaspytrthrphesaparganvalleuserarg 120
 Db 390 ATGACCGGTGTCATGTTTTCATCAGGTGATTTTCCCTTTACCGGATGTGCTATCCAG 449
 QY 121 Leuleuansglucysargluleuemetleuhsiglnleileglnarghlsleuthrallys 140
 Db 450 CTGTTGAATGAGTGACAGAGATGCTGACACAGATCATCCAGGCTCATCCACCAAG 509
 QY 141 SerHlsGlyArGyAlaenValPheasPhlsPheSerAspCySGlupheleuAla 160
 Db 510 TCACATGACGCTGTAAACAAGCTTTGATCACTTCTCAGATTGTGATTTCTTGGCCGCC 569
 QY 161 LeuTyAsnProphelglsanphelysProHlsleuGlnlyseuCyAspGlylleasn 180
 Db 570 TTGTATATATCCCTTTGAAATTTTAAACCCGACTTACAGAACTGTGTGATGTATCAAC 629
 QY 181 LysMetLeuAspGluGluAsnille 188
 Db 630 AAAATGTTGGATGAGATTAACATA 653

RESULT 22
 CA984413

LOCUS DEFINITION 942 bp mRNA linear EST 06-JAN-2003
 AGENCOURT 11295099 NIH-MGC 164 Mus musculus cDNA clone
 IMAGE:30146475 5', mRNA sequence.

ACCESSION CA984413 GI:27517069

VERSION KEYWORDS

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 942)
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC http://mgi.nci.nih.gov/.

/note="Organ: Intestine; Vector: PCMVSPORT6.1; Site_1: NCRI; Site_2: ECKR; Normalized cow cDNA intestinal mRNA library in PCMVSPORT6.1, constructed from equimolar mRNA pools derived from 5 sources, 4 lactating intestinal, 1 neonatal intestinal 4/5 lactating, proximal duodenum, jejunum, distal ileum, colon, 1/5 neonatal, proximal duodenum, jejunum, distal ileum"

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe and Dr. Mina
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM0062 row: d column: 04
High quality sequence start: 21
High quality sequence stop: 693.

FEATURES
source location/Qualifiers
1..942
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30146475"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_164"
/note="Vector: pCMV-Sport6.1, Site_1: EcoRV, Site_2: NotI;
Non-normalized full-length enriched library from pooled
mouse embryonic limb, maxilla and mandible, day 10.5 and
11.5 (size selected for the 0.5-1 kb fragments) Cloned
directionally, priming method: Oligo-dT. cDNA enrichment:
>1k bp, Average insert size 1.8k bp. Priming sequence:
5'GACTGAGTTCTAGATCGGAGCGGCGCC(T) 3'. Tissue contributed
by, David Rowe. Library constructed by Reegen, Invitrogen
Corp."

ORIGIN

Alignment Scores:
Pred. No.: 4,01e-101 Length: 942
Score: 924.00 Matches: 178
Percent Similarity: 98.94% Conservative: 8
Best Local Similarity: 94.68% Mismatches: 2
Query Match: 96.05% Indels: 0
Gaps: 0

US-10-627-571-2 (1-188) x CA984413 (1-942)

```

QY 1 MetAlAThAsPvAlPhAsnSerLySAsnLeuAlaValGlnAlaGlnLySylleu 20
Db 176 ATGGCTACGATGCTCTTCATTCCTCAAAACCTGGCCCTTCAAGCAAAAGAAAGTCTTG 235
QY 21 GilylVMeTValSerLySeriLeaIaThrThrLeuIleAspAspThrSerSergIuaI 40
Db 236 GGCAAATGATGATCCAAATCCATCCGACACGCTGATGACGACACGACGACGAGGTG 295
QY 41 LeuAspGluLeuTyArGAlThraGluTyTrThGlnAsnLySgluaIaGluLyS 60
Db 296 CTAGATGAGCTGTACGGGTGACCAAGAGTACACCAAGAAAGAGGCGGAGGAG 355
QY 61 LysIleLyAsnLeuIleLySThrValIleLySLeuAlaIleuTyArGAsnAngin 80
Db 356 GTCATCAAGAACCTCATCAAGACGTCATCAAGCTGCGCTCTCCACAGAACATCAG 415
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLySPhelYsLySValIhIsgInLeuAla 100
Db 416 TTCATCAAGACGAGCTGCGCTCATGAGAGTCAAGAGAGGTGACACGCTTGCC 475
QY 101 MetThrValValSerPhehIsgInValaSerTyTrThPhAspArGAsnValleuSerArG 120
Db 476 ATGACGGTGTGAGCTTCCACCAAGTACACCTTGAACCGCATGTGCTGTCCAG 535
QY 121 LeuLeuAsnGluCyArGluMetLeuThIsgInIleIleGlaRghIstLeuThraIaLyS 140
Db 536 CTGGCTGAACGATGCGGAGAGCTCTCAACGAGATATTCAGGCGCACTTACCGCCAG 595
QY 141 SerhIsgLyArValaAsnValPheAspHIsPheSerAspCySgluPheLeuAlaIa 160

```

```

Db 596 TCTCAGCAGCGGATTATATATGCTTTTGACCATTTTTCAGATTGATTTTGGCTGCC 655
QY 161 LeuTyArAsnProPhGlyAsnPhelYsProhIstLeuGlnLySLeuCyAspGlyIleAsn 180
Db 656 TTGTCAATCCCTTTGGAAAGTTAAACCTCATCAAGAACTTTGGACGCGCATCAAC 715
QY 181 LysMetLeuAspGluGluAsnIle 188
Db 716 AAAATGTTGGATGAAGAGACATA 739

```

RESULT 23
BI759027 893 bp mRNA linear EST 25-SEP-2001
LOCUS 603042814r1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5183355 5',
DEFINITION mRNA sequence.
ACCESSION BI759027.1 GI:15750605
VERSION BI759027.1 GI:15750605
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM11457 row: k column: 04
High quality sequence start: 5
High quality sequence stop: 869.
location/Qualifiers
1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5183355"
/lab_host="DH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-Sport6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 4.94e-101 Length: 893
Score: 923.00 Matches: 185
Percent Similarity: 98.94% Conservative: 1
Best Local Similarity: 98.40% Mismatches: 2
Query Match: 95.95% Indels: 1
Gaps: 0

US-10-627-571-2 (1-188) x BI759027 (1-893)

```

QY 1 MetAlAThAsPvAlPhAsnSerLySAsnLeuAlaValGlnAlaGlnLySylleu 20
Db 94 GTGGCCACAGATGCTTTAATTCCTCAAAACCTGGCCGTTCAAGGCAAAAGAAAGATCTTG 153

```

OY		21	GlytSmetValSerIysSerIleAlaThrTirLeuileAspAspThrsSerGluVal	40
Db		154	GCTAAATGGTGTCCTCCAAATCCATGCCACCACTTTAATAGACAACAAGTAGTAGGTG	213
OY		41	LeuAspGluLeuTyArgValThraGgJuryTrhgInAsnLysGluAgluLys	60
Db		214	CTGGATGAGCTCTACAGAGTAGCAGGGAGTACCACCAAACCAAGAGGCACAGAG	273
OY		61	LysIleIyBAenllellyeThrValilleylsLeuAlalleuTyTrghAmngln	80
Db		274	ATCATCAAGAACCCTCATAMGACAGTCATCAAGCTGGCCATTCTTAATAGGAATATCAG	333
OY		81	PheAsnGlnAspGlnLeuAlaleuMetGluLysPheLysLysValHieglInleuAla	100
Db		334	TTTTATCAAGATGACTTACCATTTATGAGGAATTTAAGAAAGAAATTCATCAGCTTCT	393
OY		101	MetThrValIalSePheHieglInValAspTyTrhrPheAspArgAsnValIeusera	120
Db		394	ATGACCGTGGTCACTGTTCCATAGGTGGATTATACCTTTGACCGGAATGTGTATCCAG	453
OY		121	LeuLeuAsnGluCyArgGluMetLeuHieglInIelleglNargHieleuthAlals	140
Db		454	CTGTTAAATGAATGCAGACAGAGCTGCACCAATCATTCAGCGCCACTCATCTCCAG	513
OY		141	SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCyGluPheLeuAlala	160
Db		514	TCCATCGGACGGCTTAATGAATGTGTTATCATTTTTCAATGTGAATTTTGG-GCTGCC	572
OY		161	LeuTyAsnPProPheGlyAsnPhelysProhisleuGlnLysLeuCyAspGlytleAsn	180
Db		573	TTGTRTATATCCTTTTGGAAATTTTAAACCCCACTTACAAAACCTATGTGATGTATCAC	632
OY		181	LysMetIeuAspGluGluAsnIle	188
Db		633	AAATGTGATGAAGAGACATA	656
RESULT_24				
LOCUS	Bf607236			
DEFINITION	Myl 0001178 Mouse 9-day fetus cDNA library ICRF522 Mus musculus	832 bp	mRNA	linear EST 01-APR-2001
ACCESSION	Bf607236			
VERSION	Bf607236.1			
KEYWORDS	EST.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE	1 (bases 1 to 832)			
AUTHORS	Yahyaoui,M., Hemmig,S., Neidhardt,L., Radeflof,U., Hermann,B.G., Lehrach,H. and O'Brien,J.			
TITLE	Detection of a high number of novel genes in a 9-day mouse embryo cDNA library normalised by oligonucleotide fingerprinting			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Hennig S Laboratory 123, dept. Leitrach Max-Planck-Institut fuer Molekulare Genetik Innesstr.63-73, D-14195 Berlin, Germany Tel.: +49 30 8413 1612 Fax: +49 30 8413 1380 Email: hennig@molgen.mpg.de EST's are made from clones being representatives of clone clusters. Clone clusters were calculated from oligonucleotide fingerprints. PCR Primers FORWARD: 5'-GAGCTATTCGAGAGTACTGA-3' BACKWARD: 5'-TAATACGACTCATTATGGG-3'. Seg primer: 5'-ATTAGTGACACTATGA-3'. High quality sequence stop: 832. Location/Qualifiers 1..832 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090"			
FEATURES				
source				

```

/clone="ICRFP522C0838"
/issue_type="whole embryo"
/dev_stage="embryonic 9-day"
/lab_host="E. coli XL1 blue"
/clone_lib="Mouse 9-day fetus cDNA library ICRFP522"
/notes="Vector: pSVSPORT1, Site 1: NotI, Site 2: SalI;
library preparation by Oligo-dT priming of RNA. Clones can
be ordered from the Resource Center in Berlin,
http://www.rzpd.de."

```

Alignment Scores:

Pred. No.:	1.04e-100	Length:	832
Score:	920.00	Matches:	177
Percent Similarity:	98.94%	Conservative:	9
Best Local Similarity:	94.15%	Mismatches:	2
Query Match:	95.63%	Indels:	0
DB:	2	Gaps:	0

US-10-627-571-2 (1-188) x BF607236 (1-832)

Qy		1	MetLathirapValPheAnsSerLYAsnLeuAlaGlnIagLnLYsLIleu	20
Db		156	GTCGCTCAAGATGCTTCCATTCCAAAACTGGCGCTTGACGCACAAAGAATCCTG	215
Qy		21	GIlyMetValSerLYSerLIEalThrThrLeuIleAspPthrSerSerGIuVal	40
Db		216	GGCAAATGGATCCAATCCATGCCACACCGGTATCGACGACACAGAGGAAGTG	275
Qy		41	LeuAspGIuLeuTYrArGVAlThrArgGLNtyrThrGlnAnLYsGLuIagLUys	60
Db		276	CTAGATAGAGCTGTAACAGGGTAGCAACAAGAGTACACCAGAACAAAGAGAGGCGAAGG	335
Qy		61	LysLIleLYsAnLeuIleLYrThrValILlELYsLeuAlalleuLYrArGVasAnGln	80
Db		336	GTCATCAAGAACTCATCAAGACGGTCATCAAGCTGGCGCTCCACAGGAACAACTAG	395
Qy		81	PheAnGlnAspGIuLeuAlaleuMetGIuLYsPheLYsLYsValHISgInLeuAla	100
Db		396	TTCATATCAAGACGAGCTGGCGCTCATAGGAAGTTCAAGAAAGAGGCGACACAGCTGGCC	455
Qy		101	MetThrValVALSerPheHISgInValAspTYrThrPheAspArgAnValLeuSerArg	120
Db		456	ATGACGGTCTCTCAAGCTTCCACACAGTGAAGTACACTTCCACCGCAATGTGCTCCAGG	515
Qy		121	LeuLeuAnGluCYsarGVImetLeuHISgInLIleIGlnArgHISLeuThralALys	140
Db		516	CTGGTGAACGAGTGGCCAGAGCTCTTACACGAGTATCTTACGGCCACTTACCGCCAG	575
Qy		141	SerHISgIYrGVAlasnAnValPheAspHISPheSerAspCYsgIuPheLeuAlALA	160
Db		576	TCTCACCGACGGGTATAATATGCTTTGACCATTTTTCAATGTGATTTTTTGGCTGCC	635
Qy		161	LeuTYrAsnProPheGIyaenPhelyrProHISLeuGlnLYsLeuCYsarGIYtIeAsn	180
Db		636	TTGTACATATCTCTTTGGAAAATTAAACCTCACTTACAGAACTTTGGCAGCGCATCAAC	695
Qy		181	LYsMetLeuAspGIuIeAnile	188
Db		696	AAATGTGATGAAGAGAACATA	719
<hr/>				
RESULT 25				
CA469176				
LOCUS		841 bp	mRNA	linear EST 09-MAR-2004
DEFINITION		AGENCOUNT_10735591 NIH_MGC_154 Mus musculus cDNA clone		
IMAGE:		30122433 5', mRNA sequence.		
ACCESSION		CA469176		
VERSION		CA469176.1 GI:24925528		
KEYWORDS		EST.		
SOURCE		Mus musculus (house mouse)		
ORGANISM		Mus musculus		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus-		

REFERENCE 1 (bases 1 to 841)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 cDNA Library Preparation: Mark Bittlinger
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: NDKM0025 row: f column: 10
 High quality sequence stop: 576.
 Location/Qualifiers
 1..841
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30122433"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Kidney; TCD (DMSO vehicle) treated 48 hours IP injections; Vector: pDONR201; Site 1: atp2; Site 2: atp1; cDNA made by oligo-dt with atp2 site and directionally cloned. Priming sequence: 5'-TTTCCGACAGCGCGCCACACTTTTACAGAAAGCTGGTTTCTTTTCTTT-3'. Full-length enriched library was constructed using the GeneRacer kit by Invitrogen, library amplification 16 cycles. Library constructed by Mark Bittlinger in the Bradfield laboratory (McArdle laboratory for Cancer Research, University of Wisconsin). Note: this is a NIH-MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 1,05e-100 Length: 841
 Score: 920.00 Matches: 177
 Percent Similarity: 98.94% Conservative: 9
 Best Local Similarity: 94.15% Mismatches: 2
 Query Match: 95.63% Indels: 0
 DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x CA469176 (1-841)
 QY 1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysIleLeu 20
 Db 184 GTGGCTACAGATGTCCTTCAATTCCAAAACCTGGCCGTTACAGCACAAAAGAAATCTTG 243
 QY 21 GlyIysMetValSerIysSerIleAlaThrThrIleuIleAspAspThrSerSerGluVal 40
 Db 244 GGCAAATGATGATCAAAATCCATCGCCACACGCTGATGACACACCGACGCGAGGTG 303
 QY 41 LeuAspGluLeuTyzArgValTharArgIuTyThGlnAsnIysGlnAlaGlnIys 60
 Db 304 CTAGATGACTGTACAGGCTGACCAAGAGTACACCCAGAAACAGAGAGGCGAGAG 363
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyzArgAsnArgin 80
 Db 364 GTCATCAAGAACCTCATCAAGAGCGTCAAGCTGCGCTCTCCACAGAAATCATG 423
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysIleLysValIleGlnLeuAla 100
 Db 424 TTCATCAAGACGAGCTGGCGCTCATGAGAGTTCAAGAGAGGTGACCAAGCTTGCC 483
 QY 101 MetThrValIleSerPheHisGlnValAlaPylTyThPheAspArgAsnValIleuSerArg 120
 Db 484 ATGACGCGTGTGACGCTTCCACAGAGTACAGTACACTTGCACCGCAATGTGCTGTCAAG 543
 QY 121 LeuLeuAsnGluCySerArgIleuMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 544 CTGCTGAACGAGTGGCGAGAGCTCTTACACAGAGATCATTCAGGCGCACCTTACCGCAAG 603

QY 141 SerHisGlyArgValIleAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 604 TCTCAGGACGGGTTAAATATGTCCTTTGACCATTTTCAGATTGATTTTGGCTGCC 663
 QY 161 LeuTyAsnProPheGlyAsnPheLysProHisLeuGlnIleLysCysAspGlyIleAsn 180
 Db 664 TTGTACATCCCTTTGGAAAGTTTAACCTCATACAGAAACTTTGGACGCAATCAAC 723
 QY 181 LysMetLeuAspGluGlnAlaIle 188
 Db 724 AAAATGTTATGATGAAGACACATA 747

RESULT 26
 B1525793 865 bp mRNA linear EST 29-AUG-2001
 LOCUS 602924756F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5057508 5',
 DEFINITION mRNA sequence.
 ACCESSION B1525793
 VERSION B1525793.1 GI:15350585
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Plate: LNA11155 row: 0 column: 13
 High quality sequence start: 28
 High quality sequence stop: 829.
 Location/Qualifiers

1..865
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C2BCH 11"
 /db_xref="taxon:10090"
 /clone="IMAGE:5057508"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NCI_CGAP_Lu33"
 /note="Organ: lung; Vector: pPT73D-Pac (pharmacia) with a modified polylinker; Site 1: NotI; Site 2: EcoRI; 1st strand cDNA was prepared from mRNA obtained from pooled lung tumors with a Not I - oligo(dt) primer (5' TGTACCAATCTGAAGTGGAGCGGCGCCCTCTCTTTTCTTTTCTTTT 3'). Double-stranded cDNA was ligated to Eco RI adaptor (pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pPT73 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES

source

ORIGIN
 Alignment Scores:
 Pred. No.: 1.09e-100 Length: 865
 Score: 920.00 Matches: 177
 Percent Similarity: 98.94% Conservative: 9
 Best Local Similarity: 94.15% Mismatches: 2
 Query Match: 95.63% Indels: 0
 DB: 4 Gaps: 0

US-10-627-571-2 (1-188) x B1525793 (1-865)
 QY 141 SerHisGlyArgValIleAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 604 TCTCAGGACGGGTTAAATATGTCCTTTGACCATTTTCAGATTGATTTTGGCTGCC 663
 QY 161 LeuTyAsnProPheGlyAsnPheLysProHisLeuGlnIleLysCysAspGlyIleAsn 180
 Db 664 TTGTACATCCCTTTGGAAAGTTTAACCTCATACAGAAACTTTGGACGCAATCAAC 723
 QY 181 LysMetLeuAspGluGlnAlaIle 188
 Db 724 AAAATGTTATGATGAAGACACATA 747

QY 1 MetAlaThrApValPheAsnSerLyAsnLeuAlaValGlnAglInlySylleu 20
 Db 95 GTGGCTACAGATGCTTCAATTCACAAACCTGGCCCTTCAGGCAAAAGAAATCTTG 154
 QY 21 G1yVemEtValSerLySerleAlaThrThreuleAspAerThrserserGluVal 40
 Db 155 GGCAGAAATGATTCACAAATTCATGCGACACCGCTGATCGACACACAGAGAGAGTG 214
 QY 41 LeuAspGluLeuValArgGluThrArgGluThrGlnAsnlySylGluAgluLys 60
 Db 215 CTAGATAGAGCTGATACAGAGTGACCAAGAGTACCCAGAACAAAGAGAGGGCGAGAG 274
 QY 61 Lye11elysAsnleuleyThrVal11elysleuAla11elyuThrArgAsnGln 80
 Db 275 GTGATCAAGAACCTTCATCAAGACGCTGATCAAGCTGGCCCTTCACAGAACAAATCAG 334
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysVal1H1sgInLeuAla 100
 Db 335 TTCATTCAGACAGCTGGCGCTCATGAGAGAGTTCCAGAGAGAGTGCCAGCTTGC 394
 QY 101 MetThrValValSerPheH1sgInValAspTyThrPheAspArgAsnValLeuSerArg 120
 Db 395 ATGACGGTCTGACGCTTCACACAGGTAGTACACCTTCGACCGCATGTGCTGTCAGG 454
 QY 121 LeuLeuAsnGluCyAsArgGluMetLeuH1sgIn11elyGlnArgH1sleuthrAlaLys 140
 Db 455 CTGCTGACAGAGTCCAGAGCTCTTCACAGATCTTCAGCGCCCTTACCGCCAG 514
 QY 141 SerH1sg1yArgValAsnAsnValPheAspH1sPheSerAspCysGluPheLeuAla 160
 Db 515 TCTCACGACGAGGTAAATGATTTGACCATTTTCAGATTGTGATTTTGGCTGCC 574
 QY 161 LeuTyAsnProPheGlyAsnPhelysProH1sleuGlnLysLeuCyAsArgLysLeu 180
 Db 575 TTGTACATCCCTTTGGAAAGTTTAAACCTCATTACAGAAACTTTCCGACGGCATCAC 634
 QY 181 LyeMetLeuAspGluLysleu 188
 Db 635 AAAATGTTGATGACAGAGACATA 658
 RESULT 27
 AL558873 559 bp mRNA linear EST 02-APR-2004
 LOCUS AL558873 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 DEFINITION Homo sapiens cDNA clone CS0DJ015H04 5-PRIME, mRNA sequence.
 ACCESSION AL558873
 VERSION AL558873.3 GI:46184260
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 L1 (bases 1 to 559) Jesse, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 15, 2001 this sequence version replaced gi:3128306.
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 3485.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0DJ015D020Qp1&c=3485.f.
 FEATURES
 source 1..559
 location/Qualifiers
 /organism="Homo sapiens"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,07e-100 Length: 559
 Score: 918.00 Matches: 180
 Percent Similarity: 98.90% Conservative: 0
 Best Local Similarity: 98.90% Mismatches: 2
 Query Match: 95.43% Indels: 0
 DB: 1 Gaps: 0
 US-10-627-571-2 (1-188) x AL558873 (1-559)
 QY 2 AlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAglInlySylleuGly 21
 Db 3 GCCACAGATGCTTAAATTCACAAACCTGGCCGTTTCAGGCAAAAGAAATCTTGAGT 62
 QY 22 LyeMetValSerLySerleAlaThrThreuleAspAerThrserserGluValLeu 41
 Db 63 AAAATGTTGATGACAGAGTCCAGAGCTCTTCACAGATCTTCAGCGCCCTTACCGCCAG 122
 QY 42 AspGluLeuValArgValThrArgGluThrArgGluThrGlnAsnlySylGluAgluLys 61
 Db 123 GATGACCTTACAGAGTACACAGAGTACACCCAAACAAAGAGAGGAGAGAGATC 182
 QY 62 11elysAsnleuleyThrVal11elysleuAla11elyuThrArgAsnGlnPhe 81
 Db 183 ATCAAGACCTTCATCAAGACAGTCAAGCTGCGCATTTTATAGGAATATCAGATT 242
 QY 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysVal1H1sgInLeuAlaMet 101
 Db 243 AATCAAGATGACCTTACATTCATGAGAGAAATTTAAAGAGAAATTCATCAGCTTGCTATG 302
 QY 102 ThrValValSerPheH1sgInValAspTyThrPheAspArgAsnValLeuSerArgLeu 121
 Db 303 ACCGTGATGATTTCCATCATGAGTATTAACCTTGACCGGATGTATTCAGGCTG 362
 QY 122 LeuAsnGluCyAsArgGluMetLeuH1sgIn11elyGlnArgH1sleuthrAlaLysSer 141
 Db 363 TTAAATGAATGACAGAGATGCTGACCAAAATCATTACAGGCGCATCTGCACAGTCA 422
 QY 142 HisG1yArgValAsnAsnValPheAspH1sPheSerAspCysGluPheLeuAla11ely 161
 Db 423 CATGACGAGTTTAAATGATTTGATTCATTTTCAGATTGTGAATTTTGGCTGCTTG 482
 QY 162 TyAsnProPheGlyAsnPhelysProH1sleuGlnLysLeuCyAsArgLysLeu 181
 Db 483 TATTAATCCCTTTTGGAAATTTTAAACCCCACTTACAAAAATATATGTATGATGATTAACAA 542
 QY 182 MetLeu 183
 Db 543 ATGTTG 548
 RESULT 28
 BP302623 580 bp mRNA linear EST 17-SEP-2004
 LOCUS BP302623 Sugano cDNA library, macrophage Homo sapiens cDNA clone
 DEFINITION MPE06324, mRNA sequence.
 ACCESSION BP302623
 VERSION BP302623.1 GI:52231583
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 580)
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.

FEATURES
 source 1..580
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MPB06324"
 /cell_type="macrophage"
 /location/Qualifiers
 /clone_lib="Sugano cDNA library, macrophage"

ORIGIN

Alignment Scores:
 Pred. No.: 1,12e-100 Length: 580
 Score: 918.00 Matches: 180
 Percent Similarity: 99.45% Conservative: 0
 Best Local Similarity: 99.45% Mismatches: 1
 Query Match: 95.43% Indels: 0
 Gaps: 0

US-10-627-571-2 (1-188) x BP302623 (1-580)

QY 8 SerLySAsnLeuAlaValGlnAglInLylsYlLeuGlyysMetValSerLySer 27
 Db 2 TCCAAAAACCTGGCCCTTCAGGACAAAGAAATCTGGGTAAATGCTCCAATCC 61
 QY 28 TleAlaThThrLeuLleAspAspThrSerSerGluValLeuAspGluLeuTyrArgVal 47
 Db 62 ATCCGACCACTTAATAGACACACAAAGTAGAGAGTGTGAGTGAAGCTTACAGAGTG 121
 QY 48 ThrArgGluTyrThrGlnAsnLysGluAglLysLysLysLysLysLysLysLys 67
 Db 122 ACCAGGAGTACACCCAAACAAAGAGGAGGACAGAGATCATCAAGAACTCATCAAG 181
 QY 68 ThrValLleLysLeuAlaLleLeuTyrArgAsnArgGlnPheAsnGlnAspGluLeuAla 87
 Db 182 ACAAGTCATCAAGCTGGCCATCTTTATAGAAATATCAAGTTATCAAGATGACTAGCA 241
 QY 88 LeuMetGluLysPheLysLysValHisGlnLeuAlaMetThrValLysPheHis 107
 Db 242 TTGATGAGAAATTTTAAAGAAAGATTCATCACTGCTGATGACCGTGTGCTTTCAT 301
 QY 108 GlnValAspTyrThrPheAspArgAsnValLleuSerArgLeuLeuAsnGluCysArgGlu 127
 Db 302 CAGGTGAGATTATACCTTTGACCGGAATGTTATTCAGGCTGTTAATGATGACAGAG 361
 QY 128 MetLeuHisGlnLleLleGlnArgHisLeuThrAlaLysSerHisGlyArgValAsnAsn 147
 Db 362 ATGCTGACCAATCATTCAGCGCCACCTCAGCCCAAGTCATCAAGAGCGGGTTAATAT 421
 QY 148 ValLleAspHisPheSerAspCysGluPheLeuAlaLleLeuTyrAsnPropheGlyAsn 167
 Db 422 GGTGTTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATCTTTTGGGAAT 481
 QY 168 PheLysProHisLeuGlnLysLeuCysAspGlyLleAsnLysMetLeuAspGluGluAsn 187
 Db 482 TTTAAACCCCATTAACAAAACATATGTGATGATCAACAAAATGTTGATGAGAGAAC 541
 QY 188 Ile 188
 Db 542 ATA 544

RESULT 29
 BGI46795 651 bp mRNA linear EST 01-FEB-2001
 LOCUS mab95h09.y1 NCI CGAP Sp2 Mus musculus cDNA clone IMAGE:397813 5
 DEFINITION similar to TR:095379 095379 MDC-3.13 ISOFORM 2. [2] TR:09UP47 ,
 mRNA sequence.

ACCESSION BGI46795.1 GI:12650203
 VERSION BGI46795
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 651)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Email: cgapbs-remail.nih.gov
 Tissue Procurement: David Segal Ph.D., Herbert Morse M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILM at:
<http://image.llnl.gov>

MGI:1478145
 Seg primer: -40RP from Gibco
 High quality sequence stop: 436.
 location/Qualifiers
 1..651
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:397813"
 /tissue_type="NK cells (flow-sorted)"
 /lab_host="DH10B (TI-resistant)"
 /clone_lib="NCI-CGAP_Sp2"
 /note="Organ: spleen; Vector: pCMV-SPORT6 (Life Technologies); mRNA made from flow-sorted NK cells, cDNA made by oligo-dT priming. Directionally cloned. Average insert size 1.5 kb. Primary library, non-amplified. cDNA library Preparation: David B. Krizman, Ph.D."

FEATURES

source

Alignment Scores:
 Pred. No.: 2,28e-100 Length: 651
 Score: 916.00 Matches: 177
 Percent Similarity: 98.40% Conservative: 8
 Best Local Similarity: 94.15% Mismatches: 3
 Query Match: 95.22% Indels: 0
 Gaps: 0

ORIGIN

US-10-627-571-2 (1-188) x BGI46795 (1-651)
 QY 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAglInLysLleLeu 20
 Db 88 ATGGCTACAGATGCTTCATTCATAAACCCTGGCCGTTACAGGCACAAAGAAATCCCG 147
 QY 21 GlyLysMetValSerLysSerLleAlaThrThrLeuLleAspAspThrSerSerGluVal 40
 Db 148 GGCAAAAGTGATCCAAATCCATCCCAACAGCTGATGACAGACCAAGCAGAGGG 207
 QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysLysGluAglLys 60
 Db 208 CTAGATGAGCTTACAGGCTGACCAAGAGATACCAAGCAAGAAAGAGCGGAGAG 267
 QY 61 LysLleLysAsnLeuLleLysThrValLleLysLeuAlaLleLeuTyrArgAsnGln 80
 Db 268 GTCATCAAGAACTCATCAAGAGCGGTCAAGCTGCGCTCCACAGAAACATCAG 327

QY 81 PheanGlnAapGluLeuA1aleuMetGluLysPheLysValHisGlnLeuA1 100
 Db 328 TTCAATCAAGACGAGCTGGCCCTCAATGAGAAAGTTCAAGAAAGAGTGACCACTGGCC 387
 QY 101 MetThrValAlaSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 388 ATGACGGTGGTCACTTCCACCAAGTAGATGACCTTCCAGCCGCAATGCTGTCCAGG 447
 QY 121 LeuLeuAanGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 448 CTGCTGAACAGAGTGGCCAGAGCTCTACACAGATCTTCAAGCCCACTTACCGCCAG 507
 QY 141 SerHisGlyArgValaAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 508 TCTCAGGAGAGGGTTAAATGATCTGTGACCATTTTCAGATTGTGATTTTGGCTGCC 567
 QY 161 LeuTyrAsnProPheGlyAsnPhelYsProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 568 TTGTACATCCCTTTGGAAGTTTAAACCTCACTTACGAAACTTTGCGACGATCAAC 627
 QY 181 LysMetLeuAapGluGluAenIle 188
 Db 628 AAAATGTTAGATGAAGAACATA 651

RESULT 30
 AUI35377 816 bp mRNA linear EST 02-AUG-2002
 LOCUS AUI35377 PLACE1 Homo sapiens cDNA clone PLACE1001920 5', mRNA
 DEFINITION
 ACCESSION AUI35377
 VERSION AUI35377.1 GI:10995916
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 816)
 Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
 HRI human cDNA project
 Unpublished (2000)
 CONTACT: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@hri.co.jp
 HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix Research Institute; cDNA library construction; Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
 FEATURES
 Source
 1. 816
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1001920"
 /tissue_type="placenta"
 /clone_id="PLACE1"
 /note="Vector: PME18SFM3"

ORIGIN
 Alignment Scores:
 Pred. No.: 3 08e-100 Length: 816
 Score: 916.00 Matches: 184
 Percent Similarity: 97.89% Conservative: 2
 Best Local Similarity: 96.84% Mismatches: 2
 Query Match: 95.22% Indels: 0
 DB: 1 Gaps: 2
 US-10-627-571-2 (1-188) x AUI35377 (1-816)

QY 1 MetAlaThrAspValPheAsnSerIyAsnLeuAlaValGlnAagLysValLeu 20
 Db 101 GTGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCAAGCACAAAAGAAATCTTG 160
 QY 21 GlyMetValSerIySerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 161 GGTAAATGTGTCCAAATTCATCGCACCACTTAAATAGACACACAGATAGAGTG 220
 QY 41 LeuAspGluLeuTyrArgValThrArgLysTyrThrGlnAsnLysGluAglLys 60
 Db 221 CTGATGAGCTTACAGAGTACAGAGGAGTACACCAAAACAAAGAGGAGGAGAG 280
 QY 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
 Db 281 ATCATCAAGAACTCATCAAGACAGATCAACCTGGCCATTTTAAATAGAAATTAAC 340
 QY 81 PheAnGlnAapGluLeuA1aleuMetGluLysPheLysValHisGlnLeuA1 100
 Db 341 TTTAATCAAGATGAGCTGATGATGAGAAATTTAAGAAAGTTCAATCAGCTTGTCT 400
 QY 101 MetThrValAlaSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 401 ATGACCGTGGTCACTTCCATCAGTGATTAACCTTGAACCGAATGTATATCCAGG 460
 QY 121 LeuLeuAanGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 461 CTGTTAAATGAATGACAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTCCAG 520
 QY 141 SerHisGlyArgValaAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 521 TCACATGAGAGGGTTAATAATGTTGTGATTTTCAGATTGTGAATTTTGGCTGCC 580
 QY 161 LeuTyrAsnProPheGlyAsnPhelYsProHisLeuGlnLysLeuCysAspGlyIleAsn 180
 Db 581 TTGTACATCCCTTTGGAAGTTTAAACCTCACTTACGAAACTTTGCGACGATCAAC 640
 QY 181 LysMet-LeuAapGluGluAenIle 188
 Db 641 AAAATGTTGATGAAGAACATA 666

RESULT 31
 AK090316 1542 bp mRNA linear HTC 03-APR-2004
 LOCUS AK090316
 DEFINITION Mus musculus 14 days embryo lung cDNA, RIKEN full-length enriched library, clone:G630049H02 product:TNF-INDUCED PROTEIN G62-1 homolog (Homo sapiens), full insert sequence.
 ACCESSION AK090316
 VERSION AK090316.1 GI:26355502
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci,P. and Hayashizaki,Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253
 MEDLINE
 PUBMED 10349636
 REFERENCES
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCES
 AUTHORS
 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM11901 row: d column: 05
 High quality sequence stop: 712.

FEATURES

SOURCE

Location/Qualifiers

1..908
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:534452"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Mam6"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
 library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Alignment Scores:

Pred. No.:	1.08e-99	Length:	908
Score:	912.00	Matches:	175
Percent Similarity:	98.40%	Conservative:	10
Best Local Similarity:	93.09%	Mismatches:	3
Query Match:	94.80%	Indels:	0
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x B1688217 (1-908)

```

QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAglNlySlySllLeu 20
DB 85 ATGGCTACAGATGCTTCAATTCACAAACCTGCCCTTCAGGCACAAAGAGATCCTG 144
QY 21 GlyMetValSerLySerIleAlaThrThrLeuIleAspThrSerSerGluVal 40
DB 145 GGCMAATGATTCATTCATTCATGCAACGCTGATTCACACACAGCAGGAGGTG 204
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySlyGlnAglNly 60
DB 205 CTAGATGAGCTGTACAGGTGACCAAGAGTACACCAAGAACAGAGGCGGAGAG 264
QY 61 LysIleLysAsnLeuIleTyThrValIleLysLeuAlaIleLeuTyArgAsnGln 80
DB 265 GTCATCAAGAACTCATCAAGAGCTGATCAAGCTGCCCTCTCCACAGAACTATCAG 324
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluTyPheLysValLysGlnLeuAla 100
DB 325 TTCAATCAAGAGCTGCGCTCATGTGAGAAAGTTCAGAGAGGTCACAGCTTGGC 364
QY 101 MetThrValValSerPheHisGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
DB 385 ATAGCGGTCTCAGCTTCCACAGGTAGATACCTTCACCCGCAATGTCTCTCAGG 444
QY 121 LeuLeuAsnGluTyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 445 CTGCTGAACAGTGCCAGAGCTCTTCACAGATCATTCAGGCGCCACTTACCGCAAG 504
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 505 TTCACAGGAGCGGTATATATGCTTTGACCATTTTCAATTTGTGATTCCTTGGCTGCC 564
QY 161 LeuTyArgAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuTyAspGlyIleAsn 180
DB 565 TTGTACATATCCCTTTGAAAGTTTAACTTACCTTACAGAACTTTTCGACGCGCTCAC 624
QY 181 LysMetLeuAspGluGlnAsnIle 188

```

Db 625 AAAATGTTGATGACAGAACCTT 648

RESULT 33

B1683693

LOCUS

DEFINITION

B1683693 766 bp mRNA linear EST 18-SEP-2001
 603306241F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5342123 5',
 mRNA sequence.

ACCESSION

B1683693

VERSION

B1683693.1 GI:15646321

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS

1 (bases 1 to 766)

TITLE

NIH-MGC <http://imgc.nci.nih.gov/>.

JOURNAL

National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT

Unpublished (1999)

CONTACT

Contact: Robert Strausberg, Ph.D.

EMAIL

Email: cgabs-r@mail.nih.gov

TISSUE

Tissue Procurement: Jeffrey Green M.D.

CDNA

CDNA Library Preparation: Life Technologies, Inc.

DNA

DNA Sequencing by: Incyte Genomics, Inc.

CLONE

Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>

PLATE

Plate: LLM11869 row: b column: 12

HIGH

High quality sequence start: 5

STOP

High quality sequence stop: 757.

LOCATION

Location/Qualifiers

FEATURES

1..766

SOURCE

/organism="Mus musculus"

MOL

/mol_type="mRNA"

STRAIN

/strain="FVB/N"

DB_XREF

/db_xref="taxon:10090"

CLONE

/clone="IMAGE:5342123"

SEX

/sex="female, virgin"

TISSUE

/tissue_type="infiltrating ductal carcinoma"

DEV

/dev_stage="5 months"

LAB

/lab_host="DH10B"

CLONE

/clone_lib="NCI_CGAP_Mam6"

NOTE

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr.
 library constructed by Life Technologies. Investigator
 providing samples: Jeffrey Green, M.D., NIH"

ORIGIN

Alignment Scores:

Pred. No.:	1.99e-99	Length:	766
Score:	909.00	Matches:	177
Percent Similarity:	98.41%	Conservative:	9
Best Local Similarity:	93.65%	Mismatches:	2
Query Match:	94.49%	Indels:	1
DB:	4	Gaps:	0

US-10-627-571-2 (1-188) x B1683693 (1-766)

```

QY 1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAglNlySlySllLeu 20
DB 142 ATGGCTACAGATGCTTCAATTCACAAACCTGCCCTTCAGGCACAAAGAGATCCTG 201
QY 21 GlyMetValSerLySerIleAlaThrThrLeuIleAspThrSerSerGluVal 40
DB 202 GGCMAATGATTCATTCATTCATGCAACGCTGATTCACACACAGCAGGAGGTG 261
QY 41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnLySlyGlnAglNly 60
DB 262 CTAGATGAGCTGTACAGGTGACCAAGAGTACACCCAGAACAGAGGCGGAGAG 121
QY 61 LysIleLysAsnLeuIleTyThrValIleLysLeuAlaIleLeuTyArgAsnGln 80

```

Db 322 GTCATCAGAACCTCATCAAGACGCTCATCAAGCTGCGCTCTCCACAGAAACATTCAG 381
 Qy PheAsnglAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 382 TTCATCAAGACGAGCTGCGCTCATGAGAACTTCAAGAGAGAGTGCACCGCTTGGC 441
 Qy 101 MetThrValIserPheHisGlnValAspTyrThrPheAspArgAnValIeuSerArg 120
 Db 442 ATGACGCTGCTGCTTCCACACGAGTAGAGTACACCTTCGACCGCATGTGCTGCAG 501
 Qy 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 502 CTGCTGAACGAGAGTCCGAGAGCTCTTACAGAGATCATTCAGCGCATCTTACCGCCAG 561
 Qy 141 SerHisGluArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
 Db 562 TTCACGAGACGGGTAAATATATGCTTTGACCACTTTTCAGATTGTATTTTGGCTGCC 621
 Qy 161 LeuTyrAsnProPheGluAsnPhelysProHisLeuGlnLysLeu-CysAspGluIleAs 180
 Db 622 TTGTACATCCCTTTGGAAAGTTTAAACCTCACTTACAGAACTTTTGGCAGCGCATCA 681
 Qy 180 nlySwetLeuAspGluGluAsnIle 188
 Db 682 CAAAATGTTGATGAAGAGAACTTA 706
 RESULT 34
 LOCUS BI103517 772 bp mRNA linear EST 26-JUN-2001
 DEFINITION 602889058F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:5044167
 ACCESSION BI103517
 VERSION BI103517.1 GI:14554410
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 772)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11121 row: C column: 16
 High quality sequence stop: 751.
 Location/Qualifiers
 1..772
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5044167"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1ib="NCI CGAP Kid14"
 /note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NciI;
 Site 2: SalI; cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.75 kb. Constructed by Life
 Technologies. Note: this is a NCI_CGAP Library."

Alignment Scores:
 Pred. No.: 3 52e-99 Length: 772
 Score: 907.00 Matches: 177
 Percent Similarity: 98.41% Conservative: 2
 Best Local Similarity: 93.65% Mismatches: 9

Query Match: 94.28% Indels: 1
 DB: 4 Gaps: 0
 US-10-627-571-2 (1-188) x BI103517 (1-772)
 Qy 1 MetAlaThrAspValPheAsnSerLysAsnLeuAlaValGlnAlGlnLysIleLeu 20
 Db 86 GTGGCTACAGATGCTTCAATTCCTCAAAACCTGGCGCTTCAGGCAAAAGAGATCTCG 145
 Qy 21 GlyMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 146 GCGAAATAGATGCCAAATCCATCGCACACGCTGATGAGACACAGACGACGAGGTG 205
 Qy 41 LeuAspGluLeuTyrArgValThrArgLysIleValGlnLysGluIleGluLys 60
 Db 206 CTAGATGAGCTGTACAGGCTGACCAAGAGTACACCAAGAACAAAGAGGCGAGAGG 265
 Qy 61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnGln 80
 Db 266 GTCATCAAGAACCTCATCAAGACGCTCATCAAGCTGCGCTCTCCACAGAACATTCAG 325
 Qy 81 PheAsnglAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 326 TTCATCAAGACGAGCTGCGCTCATGAGAGTTCAGAGAGAGTGCACCACTTGGC 385
 Qy 101 MetThrValIserPheHisGlnValAspTyrThrPheAspArgAnValIeuSerArg 120
 Db 386 ATGACGCTGCTGAGCTTCCACAGGATAGATACCTTCGACCGCATGTGCTGCCAGG 445
 Qy 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 446 CTGCTGAACGAGTGCAGAGCTCTTACAGAGATCATTCACGCGCATCTTACCGCCAG 505
 Qy 141 SerHisGluArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaIa 160
 Db 506 TTCACGAGACGGGTAAATATATGCTTTGACCACTTTTCAGATTGTATTTTGGCTGCC 565
 Qy 161 LeuTyrAsnProPheGluAsnPhelysProHisLeuGlnLysLeu-CysAspGluIleAs 180
 Db 566 TTGTACATCCCTTTGGCAAGTTTAAACCTCACTTACAGAACTTTTGGCAGCGCATCA 625
 Qy 180 nlySwetLeuAspGluGluAsnIle 188
 Db 626 CAAAATGTTGATGAAGAGAACTTA 650
 RESULT 35
 LOCUS CB991862 785 bp mRNA linear EST 01-MAY-2003
 DEFINITION AGENCOURT_13618130 NIH MGC_148 Homo sapiens cDNA clone
 ACCESSION CB991862
 VERSION CB991862
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 785)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Stefan Hanson
 CDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
 and advice from Piero Carninci (RIKEN)
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: NDAM363 row: 1 column: 23
 High quality sequence stop: 629.

FEATURES

Location/Qualifiers
 1. 785
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30337918"
 /issue_type="pre-clampic placenta"
 /lab_host="DH10B Tona"
 /clone_id="NH_MGC_148"
 /note="Organ: placenta; Vector: pBluescriptR; Site 1: all-XhoI; Site 2: BamH; Library is oligo-dT primed and directionally cloned using primer 5'-TTTTTTTTTTTNN-3', size-selected for average insert size 2.3 kb and normalized to R0.5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 3.6e-99 Length: 785
 Score: 907.00 Matches: 184
 Percent Similarity: 97.87% Conservative: 0
 Best Local Similarity: 97.87% Mismatches: 2
 Query Match: 94.28% Indels: 0
 DB: 6 Gaps: 0
 US-10-627-571-2 (1-188) x CB991862 (1-785)

QY 1 MetAaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIleValIleLeu 20
 Db 133 ATGACCAAGATGCTTAAATCCAAAACCTGGCCCTTCAGGACAAAAGAAATCTTG 192
 QY 21 GlyMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
 Db 193 GGTAAATGCTGCCAATCATCATGCCACCTTATAGCAGACAGATGAGGCT 252
 QY 41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnIleValGlnAlaGlu 60
 Db 253 CTGGATGAGCTTACAGAGTACAGGAGATACCCAAAACAAGAGGAGGACAGAG 312
 QY 61 LysIleLysAsnLeuIleTyrThrValIleLysLeuAlaIleLeuTyrArgAsnArg 80
 Db 313 ATCATCAAGAACTCATCAAGACATCATCAAGCTGCCATCTTATAGGAATTAATCAG 372
 QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
 Db 373 TTTAATCAGATGAGCTTACATGATGAGAGAAATTTAAGAAAGATTCAATCAGCTGCT 432
 QY 101 MetThrValAlaSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
 Db 433 ATGACCGTGTCTGATTCATCAGTGGATTAATACCTTACCGGATGTGTTATCCAG 492
 QY 121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
 Db 493 CTGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCCACCTCAGTGCAGAG 552
 QY 141 SerHisGluValArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
 Db 553 TCCATGAGACGGGTATTAATGCTTGTGATCATTTTCAATGTGTGAATTTTGGCTGCC 612
 QY 161 LeuTyrAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCys-AspGluIleAs 180
 Db 613 TTGTATATATCTTTTGGGATTTTAACCCACTTACAAAACATATGCTATGATCA 672
 QY 180 nLysMetLeu-AspGluGlu 186
 Db 673 CAAAATGTTGGATGAAGAG 692

RESULT 36
 AK087478

LOCUS

AK087478 1836 bp mRNA linear HTC 03-APR-2004
 Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30304C20 product:TNF-INDUCED PROTEIN GS2-1 homolog [Homo sapiens], full insert sequence.

ACCESSION

AK087478 GI:26104312

VERSION

AK087478.1

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

High-efficiency full-length cDNA cloning

TITLE

High-efficiency full-length cDNA cloning

JOURNAL

Meth. Enzymol. 303, 19-44 (1999)

MEDLINE

99279253

PUBMED

10349636

REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.

AUTHORS

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

TITLE

Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL

20499374

MEDLINE

11042159

PUBMED

11042159

REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komno, H., Akiyama, J., Nishi, K., Kitsuji, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.

AUTHORS

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichannel sequencing

TITLE

Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL

20530913

MEDLINE

11076861

PUBMED

11076861

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

AUTHORS

Functional annotation of a full-length mouse cDNA collection

TITLE

Nature 409, 685-690 (2001)

JOURNAL

5

REFERENCE

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

AUTHORS

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

TITLE

Nature 420, 563-573 (2002)

JOURNAL

6 (bases 1 to 1836)

REFERENCE

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Mizate, M., Nakamura, M., Nishi, K., Nomura, K., Numata, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

AUTHORS

Direct Submissions

TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

JOURNAL

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fanom.gsc.riken.jp/>
Location/Qualifiers

FEATURES

source

1. .1836

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:E130304C20"

/db_xref="taxon:10090"

/clone="E130304C20"

/issue_type="eyeball1"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="0 day neonate"

/note="TNF-INDUCED PROTEIN Gg2-1 homolog (Homo sapiens)

(SPTIR|Q9UP47, evidence: FASTV, 93.6%ID, 100%length,

match=565)

putative"

/note=".1818

1813. .1818

1836

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

/note="putative"

180 ntysmetleuaspjgluasnile 188
|||||
796 CAAATGTTAGTGAAGACATA 820

RESULT 37

CD701806

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euteria; Primates; Catarrhini; Homiidae; Homo.

1 (bases 1 to 622)

Liu, X.-Q., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and

Zeng, Y.-X.

Transcriptional Gene Expression Profile of Human Nasopharynx

Unpublished (2003)

Contact: Yixin Zeng

Cancer Center

Sun Yat-sen University

651 Dongfeng Road East, Guangzhou 510060, China

Tel: 86-1380-9770-743

Fax: 86-20-8775-4506

Email: yxzeng@gzsunm.edu.cn.

Location/Qualifiers

1. .622

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/issue_type="normal nasopharynx"

/clone_lib="human nasopharynx"

/note="ESTs generated from a normal nasopharynx cDNA

library from southern Chinese"

US-10-627-571-2 (1-188) x CD701806 (1-622)

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-10-627-571-2 (1-188) x CD701806 (1-622)

1 MetAlaThrAspValPheAsnSerLyAsnLeuAlaValGlnAlaGlnLySylleu 20

256 GTGGCTACAGATGCTTTCAATTCGAAACCTGGCGCTTCAGGCACAAAGAGATCTTG 315

21 GlyLysMetValSerLysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40

316 GGCNAATGGTATCCAAATCCATCCGACACACCTGATGACACACGACGAGG 375

41 LeuAspGluLeuTyrrArgValThrArgGluTyrrThrGlnAsnLySylGluValGluLys 60

376 CTAGATGAGCTGTACAGGCTGACCAAGAGTACACCAAGAAAGAGGAGGAGAG 435

61 LysIleLysAsnLeuIleLysThrValIleLys-LeuAlaIleLeuTyrrArgAsnAsnG 80

436 GTTCATCAAGACATCTCAAGACGCTCATCAAGGCTGGCGCTCTCCACGAAACATCA 495

80 nPheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysValIleGlnLeuAla 100

496 GTTCATCAAGACATCTCAAGACGCTCATCAAGGCTGGCGCTCTCCACGAAACATCA 555

100 MetThrValValSerPheHisGlnValAspTyrrThrPheAspArgAsnValleuSerAr 120

556 CAGGAGGCTGTCAAGCTTCCACAGAGTACAGTACACCTTCGACCGAATGTGCTCCAG 615

120 GLeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140

616 GTGGCTGAAGAGTGGCGAGAGCTCTTACACGAGATCTTCAGCGCCACCTTACCGCCAA 675

140 sSerHisGlyArgValAsnAsnValPheAspHisPheSerAspCyGluPheLeuAla 160

676 GTCTCAAGAGCGGCTTAATATGTCTTTGACATTTTTCGATTGATTTTGGCTCG 735

160 AleuTyrrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCyArgGlyLys 180

736 CTTGTAACAATCCCTTTGAAAGTTTAAACCTCACTTACAGAACTTTGCGACGCAATCAA 795

121 LeuLeuAsnGluCyArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140

Db 445 CTGTAATATGATTCAGAGATCTGACCAATATCTAGCCCACTTACCTGACAG 504
 Qy 141 SerHisgIyArGValAsnAenValPheAspHisPheSerAspCySgIuPheLeuAla 160
 Db 505 TCACATGACGCGGTAAATGATCTTTCATTTTCAGATGATGATTTTGGCTGCC 564
 Qy 161 LeuTyraNpRopHegIyAsnPhelyProHieLeuGlnYsLeuCyAspGlyTle 179
 Db 565 TTGTATATCTCTTTGGAAATTTTAAACCCACTTACAAAACATATGATGATGATC 621
 RESULT 38
 LOCUS B1687787 656 bp mRNA linear EST 18-SEP-2001
 DEFINITION 603314107F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5354281 5',
 mRNA sequence.
 ACCESSION B1687787
 VERSION B1687787.1 GI:15650415
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 656)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Tissue Procurement: Jeffrey Green M.D.
 Email: cgabs-remail.nih.gov
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LMNL at:
 http://image.lnl.gov
 plate: LMNL1900 row: m column: 02
 High quality sequence, scop: 632.
 Location/Qualifiers
 1..656
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5354281"
 /sex="female, virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /clone_id="NCI_CGAP_Mam6"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies, Investigator
 providing samples: Jeffrey Green, M.D., NIH"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,14e-98 Length: 656
 Score: 902.00 Matches: 173
 Percent Similarity: 98.92% Conservative: 10
 Best Local Similarity: 93.51% Mismatches: 2
 Query Match: 93.76% Indels: 0
 DB: 4 Gaps: 0
 US-10-627-571-2 (1-188) x B1687787 (1-656)
 Qy 1 MetAlaThrAspValPheAsnSerIyAsnLeuAlaValGlnAglYsIleLeu 20
 Db 87 GTGGCTACAGATGCTTCAATTCACAAAACCTGCGCTTCAGGCAAAAAGAGATCTG 146
 Qy 21 G1YlyseValSerIySerIleAlaThrThrLeuIleAspAspThrSerSergIuVal 40
 Db 147 GGCAAAATGATTCATCAATCATGCGCACCGTATTCACACACACACGAGGTG 206

Qy 41 LeuAspGluLeuTyraGValThrArgIuTyThrGlnAsnYsIleGluAglYs 60
 Db 207 CTAGATAGAGGTGTAACAGAGGTGACAGAGTATACCCAGAAAGAGAGCGGAGAG 266
 Qy 61 LyeIleYsAsnLeuIleYsThrValIleYsLeuAlaIleLeuTyraGAsnGln 80
 Db 267 GTCTACAGAACTTCATCAGACGCTCATAGCTGAGCCCTCTCCACAGAACATCAG 326
 Qy 81 PheAsnGlnAspGluLeuAlaLeuMetGluYsPheYsIySylsValIhSgIuLeuAla 100
 Db 327 TTCATCAGACAGAGCTGCGCTCATGAGAGTTCAGAAAGAAAGAGCGACCGCTGCC 386
 Qy 101 MetThrValIalSerPheHieSgIuValAspTyThrPheAspArgAsnValLeuSerArg 120
 Db 387 ATGACGGTCTGCTGAGCTTCCACAGGTGAGCTTACACCTTCACCGCATATGCTGTCAG 446
 Qy 121 LeuLeuAsnGluCyAspArgGluMetLeuHieGlnIleIleGlnArgHieLeuThrAlaYs 140
 Db 447 CTGCTGAACGAGTGCAGAGCTCTTACACAGAGTATTCAGCGCCACTTACCGCCAG 506
 Qy 141 SerHisgIyArGValAsnAenValPheAspHisPheSerAspCySgIuPheLeuAla 160
 Db 507 TCACAGACGCGGTAAATGATCTTTCAGCATTTTCAGATTTGATTTTTCGCTGCC 566
 Qy 161 LeuTyraNpRopHegIyAsnPhelyProHieLeuGlnYsLeuCyAspGlyTleAsn 180
 Db 567 TTGTATATCTCTTTGGAAAGTTTAAACCTCATTACAGAACTTTCGCGACATCAAC 626
 Qy 181 LyeMetLeuAspGlu 185
 Db 627 AAATGTGTGATGATC 641
 RESULT 39
 LOCUS CR557522 707 bp mRNA linear EST 13-JUL-2004
 DEFINITION DKFZP469E2119.F1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
 CR557522
 ACCESSION CR557522
 VERSION CR557522.1 GI:50278801
 KEYWORDS EST.
 SOURCE Pongo pygmaeus (orangutan)
 ORGANISM Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 1 (bases 1 to 707)
 Pousaka, A., Albert, R., Moosmayer, P., Schnupp, I., Wellenreuther, R.,
 Mewes, H.W., Well, B., Amid, C., Osanger, A., Fodor, G., Han, M. and
 Wiemann, S.
 Pongo pygmaeus mRNA (Pousaka, A., Albert, R., Moosmayer, P., et al.)
 Unpublished (2004)
 TITLE Pongo pygmaeus mRNA (Pousaka, A., Albert, R., Moosmayer, P., et al.)
 JOURNAL Contact: MIPS
 COMMENT Contact: MIPS
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,14e-98 Length: 707
 Score: 902.00 Matches: 173
 Percent Similarity: 98.92% Conservative: 10
 Best Local Similarity: 93.51% Mismatches: 2
 Query Match: 93.76% Indels: 0
 DB: 4 Gaps: 0
 US-10-627-571-2 (1-188) x B1687787 (1-656)
 Qy 1 MetAlaThrAspValPheAsnSerIyAsnLeuAlaValGlnAglYsIleLeu 20
 Db 87 GTGGCTACAGATGCTTCAATTCACAAAACCTGCGCTTCAGGCAAAAAGAGATCTG 146
 Qy 21 G1YlyseValSerIySerIleAlaThrThrLeuIleAspAspThrSerSergIuVal 40
 Db 147 GGCAAAATGATTCATCAATCATGCGCACCGTATTCACACACACACGAGGTG 206

Alignment Scores:

Pred. No.:	1,67e-98	Length:	707
Score:	901.00	Matches:	176
Percent Similarity:	99.44%	Conservative:	2
Best Local Similarity:	98.32%	Mismatches:	1
Query Match:	93.66%	Indels:	0
DB:	7	Gaps:	0

US-10-627-571-2 (1-188) x CRS57522 (1-707)

```

QY      1 MetAlaThrAspValPheAsnSerIlyAsnLeuAlaValGlnAlaGlnIlySylleu 20
      170 GTGGCCACAGATGTCTTAAATCCAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTG 229
      21 GlyIyMetValSerIlySerIleAlaThrIleuIleAspThrSerSerGluVal 40
      230 GGTAATAATGGTGTCCAAATCCATGCGCACCACTTAATAGACACAGAGTACTGAGGTG 289
      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnIlySylleuAlaGluIly 60
      290 CTGGATGAGCTCTACAGAGTGCACAGAGTACATCTCAAAACAGAGAGGAGGCAGAGAG 349
      61 IlyIleIlyAsnLeuIleIlySerIleAlaThrIleuIleAspThrSerSerGluVal 80
      350 ATCATCAAGAACTCATCAAGACAGATCATCAAGCTGGCCATCTTATAGATTAATCAG 409
      81 PheAsnGlnAspGluLeuAlaLeuMetGluIlyPheIlySylleuValIleGlnLeuAla 100
      410 TTAAATCAAGATGAGTATGATGAGGAAATTTAAAGAAAGATTCATCAGCTTGCT 469
      101 MetThrValAlaSerPheIleGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
      470 ATGACCGTGTGCTCATGTTCCATCAGGTGATTAATCTTTGACCGAATGTGTATCCAG 529
      121 LeuLeuAsnGluCysArgGluMetLeuIleGlnIleIleGlnArgHisLeuThrAlaIys 140
      530 CTGTTAAATGAATGACAGAGAGATGCTGCACCAATCATTCAGGTCACCTCACTGCCAAG 589
      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerArgGluPheLeuAlaIle 160
      590 TCACATGAGACGGGTAAATATGCTTGTGATCATTTTTCAGATTGTGATTTTGGCTGCC 649
      161 LeuTyAsnProPheGlyAsnPheIlyProHisLeuGlnIlyLeuCysAspGlyIle 179
      650 TTGTAATATCTTTTGGGAATTTTAAACCCCACTTCAAAAACCTAATGTGATGTATC 706

```

RESULT 40
CVO23086 566 bp mRNA linear EST 20-AUG-2004
LOCUS 5463 Full Length cDNA from the Mammalian Gene Collection Homo
DEFINITION sapiens cDNA 5' similar to BC007014, mRNA sequence.

ACCESSION CVO23086
VERSION CVO23086.1 GI:51480836
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 566)
Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,
Driscoll,A., Li,N., Rosenberg,J., Lamesch,P., Vidale,J., P.O.,
Clingensmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,
Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C.,
Vandenhaute,J., Cusick,M.B., Alcala,J.S., Hill,D.E. and Vidal,M.
Human ORFome Version 1.1: a Platform for Reverse Proteomics
Genome Res. (2004) In press
Contact: Vidal M
Manc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739

Email: Marc.Vidal@fci.harvard.edu
ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
results from a PCR reaction using an MGC full-length cDNA as
template DNA and ORF specific primers
PCR Primers
FORWARD: ATGCATCTCGAAGACAGAAATCCAA
BACKWARD: CATATGTTCTCTCATCCACATTT
Insert length: 566 Std Error: 21.00
Plate: 11075 row: 05 column: F
Seq primer: ACTGGCGTCTTACACGTCGTGACTGGGAAAC
High quality sequence start: 104
High quality sequence stop: 565
POLYA=No.

FEATURES

source

location/Qualifiers

```

1..566
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/feature_type="mixed"
/clone_lib="Full Length cDNA from the Mammalian Gene
Collection"
/note="Vector: mixed. The ORFs were PCR amplified from the
MGC (Mammalian Gene Collection) as of April 2004 and
cloned by recombinational Gateway cloning into pDONR223
donor vector. Reference: MGC (Mammalian Gene Collection)
Program Team, Generation and Initial Analysis of more than
15,000 Full-Length Human and Mouse cDNA Sequences. PNAS,
2002, 99(126), 16899-16903"

```

ORIGIN

Alignment Scores:

Pred. No.:	1,64e-98	Length:	566
Score:	900.00	Matches:	176
Percent Similarity:	99.44%	Conservative:	1
Best Local Similarity:	98.88%	Mismatches:	1
Query Match:	93.56%	Indels:	0
DB:	7	Gaps:	0

US-10-627-571-2 (1-188) x CVO23086 (1-566)

```

QY      1 MetAlaThrAspValPheAsnSerIlyAsnLeuAlaValGlnAlaGlnIlySylleu 20
      31 GTGGCCACAGATGTCTTAAATCCAAAACCTGGCGCTTCAGGCACAAAAGAGATCTTG 90
      21 GlyIyMetValSerIlySerIleAlaThrIleuIleAspThrSerSerGluVal 40
      91 GGTAATAATGGTGTCCAAATCCATGCGCACCACTTAATAGACACAAAGTATGAGGTG 150
      41 LeuAspGluLeuTyArgValThrArgGluTyThrGlnAsnIlySylleuAlaGluIly 60
      151 CTGGATGAGCTCTACAGAGTGCACAGAGTACATCTCAAAACAGAGAGGAGGCAGAGAG 210
      61 IlyIleIlyAsnLeuIleIlySerIleAlaThrIleuIleAspThrSerSerGluVal 80
      211 ATCATCAAGAACTCATCAAGACAGATCATCAAGCTGGCCATCTTATAGATTAATCAG 270
      81 PheAsnGlnAspGluLeuAlaLeuMetGluIlyPheIlySylleuValIleGlnLeuAla 100
      271 TTAAATCAAGATGAGTATGATGAGGAAATTTAAAGAAAGATTCATCAGCTTGCT 330
      101 MetThrValAlaSerPheIleGlnValAspTyThrPheAspArgAsnValLeuSerArg 120
      331 ATGACCGTGTGCTCATGTTCCATCAGGTGATTAATCTTTGACCGAATGTGTATCCAG 390
      121 LeuLeuAsnGluCysArgGluMetLeuIleGlnIleIleGlnArgHisLeuThrAlaIys 140
      391 CTGTTAAATGAATGACAGAGAGTGTGCACCAATCATTCAGCGCCACCTCCTCCAG 450
      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerArgGluPheLeuAlaIle 160
      451 TCACATGAGACGGGTAAATATGCTTGTGATCATTTTTCAGATTGTGAAATTTTGGCTGCC 510
      161 LeuTyAsnProPheGlyAsnPheIlyProHisLeuGlnIlyLeuCysAspGlyIle 178

```

Db 511 TTGATATATCTTTGGGAATTTAAACCCACCTTACAAACCTATGATGCT 564

RESULT 41
AL542926 779 bp mRNA linear EST 24-MAR-2004
LOCUS AL542926 Homo sapiens PLACENTA Homo sapiens cDNA clone CS00E013Y017
DEFINITION 5-PRIME, mRNA sequence.
ACCESSION AL542926
VERSION AL542926.3 GI:45718495
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 779)
AUTHORS Li, W. B., Gruber, C., Jeessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:30548565.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremlieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 3485.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdnas=CS00E013CE09QPl&c=3485.f.
Location/Qualifiers
1. 779
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00E013Y017"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Alignment Scores:

Pred. No.: 4.38e-98 Length: 779
Score: 898.00 Matches: 178
Percent Similarity: 95.19% Conservative: 0
Best Local Similarity: 95.19% Mismatches: 0
Query Match: 93.35% Indels: 0
DB: 1 Gaps: 0

US-10-627-571-2 (1-188) x AL542926 (1-779)

Qy 2 AAThtraspValPheanSerLysAsnLeuAlaValGlnAlaGlnLysLysIleLeuGly 21
Db 3 GCCACAGATGCTCTTAATTCACAAAACCTGCGCTTACGACCAAAAAGATCTTGCTG 62
Qy 22 LysMetValSerLysSerIleLeuAlaThrThrLeuIleAspAspThrSerSerGluValLeu 41
Db 63 AAAAMGCTGTCCAAATCATCGCCACCACTTAAMARAMGACCAAGTAGAGTGCTG 122
Qy 42 AspGluLeuThrValThrArgGluThrThrGlnAsnLysLysGluAlaGluLysLys 61
Db 123 GATGAGCTCTTCARAGTGAACGAGGAGTACACCCAAACAAAGAGGAGCAAGAAATC 182
Qy 62 TleAsnLeuIleLeuThrValIleLeuLeuAlaIleLeuThrArgAsnAsnGlnPhe 81
Db 183 ATCAAGAACTCATCAAGACGATCATCAAGCTGCGCATCTTTATAGAAATAMCAGTTT 242

FEATURES

source

Qy 82 AsnGlnAspGluLeuAlaLeuMetGluLysPheLysValHisGlnLeuAlaMet 101
Db 243 AATCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 302
Qy 102 ThrValValSerPheHisGlnValAspThrThrPheAspArgAsnValLeuSerArgLeu 121
Db 303 ACCGTGTCAGTTCATTCATGATGATGATGATGATGATGATGATGATGATGATG 362
Qy 122 LeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLysSer 141
Db 363 TTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 422
Qy 142 HisGluArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaLeu 161
Db 423 CATGACGGGTAT 482
Qy 162 TTAAsnProPheGluAsnPheLysProHisLeuGlnLysLeuCysAspGluIleAsnLys 181
Db 483 TATATATCTTTTGGGAATTTTAAACCCACCTTACAAACCTATGATGATGATGAT 542
Qy 182 MetLeuAspGluGluLeuLeu 188
Db 543 ATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 563

RESULT 42
CD520304 747 bp mRNA linear EST 06-JUN-2003
LOCUS CD520304
DEFINITION IMAGE:30410510 5', mRNA sequence.
ACCESSION CD520304
VERSION CD520304.1 GI:31452022
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 747)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-remail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: NDCM200 row: m column: 15
High quality sequence stop: 499.
Location/Qualifiers
1. 747
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30410510"
/tissue_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 191"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcatggcc);
Site 2: SfiI (ggcgccctggcc); Library is oligo-dT primed
and directionally cloned. PBMC Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulation
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGGCCATATGACC-3' and 3' adaptor sequence:
5'-ATCTGAGGCGGAGCGGCCGACGATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0.kb). 15/15 colonies contained inserts

FEATURES

source

ORIGIN

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

Alignment Scores:

Pred. No.:	1,67e-97	Length:	747
Score:	893.00	Matches:	176
Percent Similarity:	97.27%	Conservative:	2
Best Local Similarity:	96.17%	Mismatches:	5
Query Match:	92.83%	Indels:	0
DB:	6	Gaps:	0

US-10-627-571-2 (1-188) x CD520304 (1-747)

QY	1	MetAlThAspValpheaSenSerlyAsnLeuAlaValGlnAaglnlyslleu	20
DB	106	GTGGCCACAGATCTTAAATCCAAAACCTGCGCTCAGCAGCAAAAGATCTTG	165
QY	21	GLYMeValSerlySerllealathThleuileaspaSPThrSerSerGluVal	40
DB	166	GGTAAATGGTGTCCAAATCCATCGCCACCACTTATNAGACACAACTAGTGAGTG	225
QY	41	LeuAsGluLeuTyArgValThraGluTyThGlnaenlyslsGluAlaGluLy	60
DB	226	CTGGATGAGCTCTACAGAGTACAGGAGTACACCAAAACAAAGAGGAGAGAG	285
QY	61	LyAlleAsnleuileuThraValleleleleleleleleuTyArgaAsnGln	80
DB	286	ATCATCAAGACCTCATCAAGACGTCATACCTGCGCTTATNAGAAATATCAG	345
QY	81	pheaSnglnAspGluLeuAlaLeuMetGluLyPheLySllyValInsglnLeuAla	100
DB	346	TTTAAATCAAGATGAGTATGATGAGGAGAAATTTAAGAAAGTTCATCGCTTGT	405
QY	101	MetThValValSerPheHlsglnValAspTyThPheAspArgAsnValleuSerArg	120
DB	406	ATACCGTGGTCTCATGCTTCATCAGTGGATTAATACCTTTCACCGAATGTGATCAGG	465
QY	121	LeuLeuAsnGluCyAspArgGluMetLeuHlsglnlelelelelelelelelelelele	140
DB	466	CTGTAAAGATGATGAGAGAGAGAGTGCACCAATCTTAAGGCCACCTCATCTGCAGG	525
QY	141	SeThlsglYArgValAsnAsnValPheAspHlPheSerAspCysGluPheLeuAla	160
DB	526	TCACATGAGACGGGTATATATATGATGATGATGATGATGATGATGATGATGATGATG	585
QY	161	LeuTyArgAsnProPheGlyAsnPhelysProHlsglnLyLeuCyAspGlyleuAsn	180
DB	586	TGTATTAATCCCTTTTGGGAATTTTAACCCCACTTCAAAAACTAATGTGATGAGATCA	645
QY	181	LyMetLeu 183	
DB	646	CAATATTTG 654	

RESULT 43 BY743355 LOCUS BY743355 DERINITION BY743355 RIKEN full-length enriched, bone marrow macrophage Mus musculus cDNA clone I83007M06 5', mRNA sequence. ACCESSION BY743355 VERSION BY743355.1 GI:27169364 KEYWORDS EST. SOURCE Mus musculus (house mouse) ORGANISM Mus musculus (house mouse) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 671) Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, T., Bono, H., Kondo, S., Nikaide, I., Otsu, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schumacher, C., Gojodori, T., Balderelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H.,

TITLE

JOURNAL MEDLINE PUBLISHED

COMMENT

Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedziarski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Perce, G., Pele, G., Petrovsky, N., Pillai, R., Pontius, J.U., Q.D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semp, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wymshew-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavoian, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Aizawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterson, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resesc.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Aizawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M., and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichipillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by David A. Hume (Dept. of Biochemistry and Microbiology/Parasitology Institute for Molecular Bioscience University of Queensland Brisbane, Q 4072 Australia) whose assistance we gratefully acknowledge.

Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

Location/Qualifiers

1..671

/organism="Mus musculus"

/mol_type="RNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="I83007M06"

/tissue_type="bone marrow"

FEATURES

source

ORIGIN /cell_type="macrophage"
/clone_1ib="RIXEN full-length enriched, bone marrow
macrophage"

Alignment Scores:

Pred. No.: 1,91e-97 Length: 671
Score: 892.00 Matches: 172
Percent Similarity: 97.31% Conservative: 9
Best Local Similarity: 92.47% Mismatches: 5
Query Match: 92.72% Indels: 0
DB: 6 Gaps: 0

US-10-627-571-2 (1-188) x BF743355 (1-671)

```

QY 1 MetAlaThrAspValPheAsnSerIlyAsnLeuAlValGlnAlaGlnIlySylIleu 20
DB 110 GTGGCTACAGATGCTTCAATTCCAAAAACCTGGCCGTCAGGCACAAAAGATCTCG 169
QY 21 GlyValMetValSerIlySerIleAlaThrIleuAlaAspAspThrSerSerGluVal 40
DB 170 GGCMAATGGTATCCAAATCCATGCCACCCATGTCAGACACACGACGAGGTG 229
QY 41 LeuAspGluLeuThrArgValThrArgGluThrGlnAsnIlySylGlnAlaGluVal 60
DB 230 CTGATGAGCTGTACAGGCTGACCAAGAGTACACCCAGAACAGAGGCGGAGAGG 289
QY 61 LysIleLysAsnLeuIleIlySerThrValIleLysLeuAlaIleLeuThrArgAsnAngin 80
DB 290 GTGATCAAGAACCTCATCAAGACGCTCATCAAGCTGGCCGCTCCACAGGAACATCG 349
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 350 TTCATCAACAGACGAGCTGGCTCATGAGAGAGTCAAGATTAAGGTGACACGCTTGC 409
QY 101 MetThrValIleSerPheHisGlnValAspIlyThrPheAspArgAsnValLeuSerArg 120
DB 410 ATGACGGTGTGCTGCTTCCACCGATGAGATACCTTCGACCGCAATGCTGTCCAG 469
QY 121 LeuLeuAsnGluLysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 470 CTGCTCAACGAGTGGCGGAGGCTCTACACAGATCATTCAGCGCACCTTACCGCAAG 529
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 530 TCTCAGGAGCGGCTTATATGCTTTCGACCAATTTTCAGATTGTGATTMTTGGCTGCC 589
QY 161 LeuIlyAsnProPheGlyAsnPheLysProHisLeuGlnIlyLeuLysCysAspGlyIleAsn 180
DB 590 TTGTACAAATCCCTTTGGAAAGTTTAAACCTCACTTACAGAACTTTGGCAGCGCATCAAC 649
QY 181 LysMetLeuAspGluGlu 186
DB 650 AATAATGTTAGTCAAGAG 667

```

RESULT 44

BF140519 739 bp mRNA linear EST 24-OCT-2000
LOCUS 601787536F1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:4015160 5',
DEFINITION mRNA sequence.

ACCESSION BF140519
VERSION BF140519.1 GI:10979546
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 739)

REFERENCE NIH-MGC <http://mgi.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov

FEATURES

source

Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Inceye Genomics, Inc.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.jnl.gov>
Plate: LLM9260 row: h column: 09
High quality sequence stop: 709.
Location/Qualifiers

1..739

/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH 11"
/db_xref="taxon:10090"
/clone="IMAGE:4015160"
/tissue_type="cumor, metastatic to mammary"
/lab_host="DH10B"
/clone_1ib="NCI_CGAP_Lu30"
/note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NciI;
Site_2: SalI; transgenic model WNT-1, expression driven by
MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo
dT. Library constructed by Life Technologies.
Investigator providing samples: Gilbert Smith, NIH"

ORIGIN

Alignment Scores:

Pred. No.: 2.88e-97 Length: 739
Score: 891.00 Matches: 176
Percent Similarity: 98.40% Conservative: 9
Best Local Similarity: 93.62% Mismatches: 3
Query Match: 92.62% Indels: 1
DB: 2 Gaps: 0

US-10-627-571-2 (1-188) x BF140519 (1-739)

```

QY 1 MetAlaThrAspValPheAsnSerIlyAsnLeuAlValGlnAlaGlnIlySylIleu 20
DB 100 GTGGCTACAGATGCTTCAATTCCAAAAACCTGGCCGTCAGGCACAAAAGATCTCG 159
QY 21 GlyValMetValSerIlySerIleAlaThrIleuAlaAspAspThrSerSerGluVal 40
DB 160 GGCMAATGGTATCCAAATCCATGCCACCGCTCATGAGACACCGACGAGGTG 219
QY 41 LeuAspGluLeuThrArgValThrArgGluThrGlnAsnIlySylGlnAlaGluVal 60
DB 220 CTGATGAGCTGTACAGGCTGACCAAGAGTACACCGAACAGAGAGCGGAGAG 279
QY 61 LysIleLysAsnLeuIleIlySerThrValIleLysLeuAlaIleLeuThrArgAsnAngin 80
DB 280 GTCATCAAGAACCTCATCAAGACGCTCATCAAGCTGGCCGCTCCACAGAAACATCAG 339
QY 81 PheAsnGlnAspGluLeuAlaLeuMetGluLysPheLysLysValHisGlnLeuAla 100
DB 340 TTCATCAAGACGAGCTGGCTCATGAGAGTTCAGAAAGAGTGCACAGCTTGC 399
QY 101 MetThrValIleSerPheHisGlnValAspIlyThrPheAspArgAsnValLeuSerArg 120
DB 400 ATGACGGTGTGCTGCTTCCACAGGTAGATACCTTGCAGCGCACTGCTTCCAG 459
QY 121 LeuLeuAsnGluLysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
DB 460 CTGCTCAACGAGTGGCGGAGAGCTCTACACGAGATCATTCAGCGCACCTTACCGCAAG 519
QY 141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAla 160
DB 520 TCTCAGGAGCGGCTTATATGCTTTCGACCAATTTTCAGATTGTGATTMTTGGCTGCC 578
QY 161 LeuIlyAsnProPheGlyAsnPheLysProHisLeuGlnIlyLeuLysCysAspGlyIleAsn 180
DB 579 TTGTACAAATCCCTTTGGAAAGTTTAAACCTCACTTACAGAACTTTGGCAGCGCATCAAC 638
QY 181 LysMetLeuAspGluGluAsnIle 188

```

```

Db      639 AAAATGTCGATGACGAGAAACATA 662
|||||
RESULT 45
CR771358
LOCUS   DKEFZp46910535_r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone
DEFINITION
CR771358
ACCESSION
CR771358
KEYWORDS
SOURCE  CR771358.1 GI:52614631
EST.
Pongo pygmaeus (orangutan)
ORGANISM
Pongo pygmaeus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pongo.
1 (bases 1 to 692)
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Weilenreuther,R.,
Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
Pongo pygmaeus mRNA (Poustka,A., Albert,R., Moosmayer,P., et al.)
JOURNAL
Unpublished (2004)
COMMENT
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email s.wiemann@dkfz-heidelberg.de; any. Please contact RZPD for
ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKEFZp46910535
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.
FEATURES
source
1..692
location/Qualifiers
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKEFZp46910535"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_idb="469 (synonym: pkid1)"
/note="Vector: pSport1_sfi; Site_1: sfi1a; Site_2: sfi1b"

```

ORIGIN

Alignment Scores:

```

Pred. No.:      4,6e-97      Length:      692
Score:          889.00      Matches:      178
Percent Similarity: 98.90%      Conservative: 2
Best Local Similarity: 97.80%      Mismatches: 2
Query Match:     92.41%      Indels:      1
DB:              7          Gaps:          0

```

US-10-627-571-2 (1-188) x CR771358 (1-692)

```

QY      1 MetAlaThrAspValPheAsnSerIysAsnLeuAlaValGlnAlaGlnIysLysIleLeu 20
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      148 GTGGCCACAGATGTCCTTAATTCGAAACCTGCGCCGTCACGCAAAAGAAATCTTG 207
QY      21 GlyLysMetValSerIysSerIleAlaThrThrLeuIleAspAspThrSerSerGluVal 40
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      208 GGTAAATGTGTGTCGAAATCCATCGCCACCACTTAATAGACGACAGAGTAGTGAGTG 267
QY      41 LeuAspGluLeuTyrArgValThrArgGluTyrThrGlnAsnLysGlnAlaGlnLys 60
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      268 CTGGATGAGCTCTCAAGAGTGAACGAGGAGTACATCAAAACAAAGAGGAGGAGGAAG 327
QY      61 LysIleLysAsnLeuIleLysThrValIleLysLeuAlaIleLeuTyrArgAsnAsnGln 80
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      328 ATCATCAAGAACCTCATCAAGACAGTCATCAAGCTGCGCATTTCTTAAGAAATATCAAG 387
QY      81 PheAsnGlnAspGluLeuAlaLeuMetGlnLysPheLysLysValHisGlnLeuAla 100
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      388 TTAAATCAAGATGAGCTACATTCATGATGAGAAATTTAAGAAAGAAAGTTCATCAGCTTGCT 447

```

```

QY      101 MetThrValValSerPheHisGlnValAspTyrThrPheAspArgAsnValLeuSerArg 120
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      448 ATGACCGTGGTCAGTTTCATCAGGTGGATTAATACCTTGGACCGAAGATGTATTCACAG 507
QY      121 LeuLeuAsnGluCysArgGluMetLeuHisGlnIleIleGlnArgHisLeuThrAlaLys 140
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      508 CTGTTAAATGANTGACGAGAGATGCTGCACCAATCATTCAGCGCTCACCTCAGTCCAG 567
QY      141 SerHisGlyArgValAsnAsnValPheAspHisPheSerAspCysGluPheLeuAlaAla 160
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      568 TCACATGACGCGGTTAAATPATGTCTTGTGATCATTTTTCAGATTGTGATTTTGGCTGCG 627
QY      161 LeuTyrAsnProPheGlyAsnPheLysProHisLeuGlnLysLeuCysAspGlyIleLeu 180
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      628 TTGTAAATCTTTTGGGAAATTTTAAACCCCACTTACA-AAACTAATGTGATGTATCAAC 686
QY      181 LysMet 182
      ::::::::::
Db      687 AAAATG 692

```

Search completed: July 28, 2005, 21:30:47
 Job time : 3024 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: July 28, 2005, 15:10:58 ; Search time 3088 Seconds
(without alignments)
9430.565 Million cell updates/sec

Title: US-10-627-571-1_COPY_100_700

Perfect score: 601

Sequence: 1 cctcagcgccgcgcgcgt.....tggatgaagaacatatga 601

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_ncg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	588.4	97.9	1892	9	AF070671	AF070671 Homo sapi
2	588.2	97.9	1943	6	BD153536	BD153536 110 human
3	586.8	97.6	1921	6	BD156880	BD156880 Primer fo
4	586.8	97.6	1921	6	AX877846	AX877846 Sequence
5	586.8	97.6	1921	6	AK001931	AK001931 Homo sapi
6	584.4	93.9	1729	6	BD156785	BD156785 Primer fo
7	584.4	93.9	1729	6	AX877673	AX877673 Sequence
8	584.4	93.9	1729	9	AK001850	AK001850 Homo sapi
9	584.4	93.9	1814	9	BC005352	BC005352 Homo sapi
10	584.4	93.9	1915	9	AF099935	AF099935 Homo sapi
11	584.4	93.9	1988	6	C0726075	C0726075 Sequence
12	584.4	93.9	2003	9	AF099936	AF099936 Homo sapi
13	584.4	93.9	2081	9	BC007014	BC007014 Homo sapi
14	584.4	93.9	2502	9	AK097284	AK097284 Homo sapi
15	584.4	93.9	112626	9	AC035144	AC035144 Homo sapi
16	584.4	93.9	156277	9	AC026795	AC026795 Homo sapi
17	584.4	93.9	158057	9	AC138612	AC138612 Homo sapi
18	581.2	93.4	573	9	CR457137	CR457137 Homo sapi
19	547.6	91.1	816	6	BD149395	BD149395 Primer fo

20	547.6	91.1	816	6	AX869333	AX869333 Sequence
21	500.8	83.3	1766	9	AK097884	AK097884 Homo sapi
22	454.2	75.6	237561	2	AC095257	AC095257 Rattus no
23	444.4	73.9	1728	10	BC009090	BC009090 Mus muscu
24	444.4	73.9	184327	10	AC120859	AC120859 Mus muscu
25	408.6	68.0	2009	5	AJ720906	AJ720906 Gallus ga
26	394.6	65.7	587	6	BD149633	BD149633 primer fo
27	394.6	65.7	587	6	AX869571	AX869571 Sequence
28	355.6	59.2	1559	5	CR760636	CR760636 Xenopus t
29	354	58.9	1193	5	BC072904	BC072904 Xenopus l
30	304	50.6	515	9	AF098933	AF098933 Homo sapi
31	257.4	42.8	164258	2	AC118938	AC118938 Mus muscu
32	257.4	42.8	195574	2	AC107848	AC107848 Mus muscu
33	256.4	42.7	3986	5	BC052765	BC052765 Danio rer
34	256.4	42.7	205949	5	BX649252	BX649252 Zebrafish
35	255.8	42.6	188937	2	AC012678	AC012678 Homo sapi
36	255.8	42.6	189796	9	AC073964	AC073964 Homo sapi
37	254.4	42.3	1248	6	C0736656	C0736656 Sequence
38	254.4	42.3	1986	6	C0841622	C0841622 Sequence
39	254.4	42.3	1986	9	AK123281	AK123281 Homo sapi
40	252.4	42.0	1589	5	AJ720336	AJ720336 Gallus ga
41	250.2	41.6	110000	2	AC115187_1	Continuation (2 of
42	250.2	41.6	256285	2	AC115505	AC115505 Rattus no
43	250.2	41.6	273325	2	AC129440	AC129440 Rattus no
44	231.2	38.5	1938	5	BC076797	BC076797 Xenopus l
45	225.4	37.5	1813	5	BC053167	BC053167 Danio rer
46	225.4	37.5	189797	5	BX927313	BX927313 Zebrafish
47	220.6	36.7	134497	2	CR407586	CR407586 Danio rer
48	213.8	35.6	2175	5	BC061657	BC061657 Xenopus l
49	209.6	34.9	1357	5	BC053238	BC053238 Danio rer
50	207.8	34.6	32360	9	AC005339	AC005339 Homo sapi
51	204.2	34.0	2209	9	BC017672	BC017672 Homo sapi
52	203.4	33.8	107568	2	AC143988	AC143988 Rattus no
53	203.4	33.8	230310	2	AC120079	AC120079 Rattus no
54	199.8	33.2	2307	10	BC032199	BC032199 Mus muscu
55	199.8	33.2	180238	2	AC073688	AC073688 Mus muscu
56	199.8	33.2	222610	10	AC026385	AC026385 Mus muscu
57	199.8	33.2	232190	2	AC074167	AC074167 Mus muscu
58	199.6	33.2	561	6	C0732643	C0732643 Sequence
59	195.4	32.5	149954	2	AC151645	AC151645 Dasyatis n
60	182	30.3	1200	9	BC063014	BC063014 Homo sapi
61	182	30.3	1268	6	AX078255	AX078255 Sequence
62	182	30.3	3116	9	HSM804686	HSM804686 Homo sapi
63	182	30.3	154526	2	AL592111	AL592111 Homo sapi
64	182	30.3	159148	9	AL592424	AL592424 Human DNA
65	180.4	30.0	1196	9	AK027120	AK027120 Homo sapi
66	180.4	30.0	1433	10	BC079019	BC079019 Rattus no
67	180.4	30.0	174234	10	AC117098	AC117098 Rattus no
68	179.4	29.9	1198	10	BC058679	BC058679 Mus muscu
69	179.4	29.9	148675	10	AC131769	AC131769 Mus muscu
70	179.4	29.9	205702	10	AC140190	AC140190 Mus muscu
71	179.4	29.9	206230	10	AC084272	AC084272 Mus muscu
72	178.8	29.8	207952	2	AC151020	AC151020 Callithrix
73	176	29.3	555	5	CR457375	CR457375 Homo sapi
74	174	29.0	181230	2	AC149856	AC149856 Papio anu
75	174	29.0	182245	2	AC149842	AC149842 Papio anu
76	171.6	28.6	1171	9	AF271774	AF271774 Homo sapi
77	170.4	28.4	208275	2	AC147191	AC147191 Oculomur
78	165	27.5	217685	2	AC149677	AC149677 Bos tauru
79	159.2	26.5	340	6	AX898564	AX898564 Sequence
80	159.2	26.5	340	6	BD034097	BD034097 Sequence
81	154.6	25.7	2156	3	AK112519	AK112519 Clona int
82	146.2	24.3	829	6	C0580257	C0580257 Sequence
83	146.2	24.3	1821	3	AY095033	AY095033 Drosophil
84	126.4	21.0	252	6	AX898571	AX898571 Sequence
85	126.4	21.0	252	6	BD034104	BD034104 Sequence
86	116	19.3	4406	6	C0580256	C0580256 Sequence
87	116	19.3	69208	2	AC020466	AC020466 Drosophil
88	116	19.3	175118	3	AC010842	AC010842 Drosophil
89	116	19.3	188272	3	AC005639	AC005639 Drosophil
90	116	19.3	295225	3	AE003461	AE003461 Drosophil
91	85.4	14.2	714	6	BD146713	BD146713 Primer fo
92	85.4	14.2	714	6	AX866651	AX866651 Sequence

93	85.4	14.2	1602	6	BD160707	Primer fo
94	85.4	14.2	1602	6	AX884081	Sequence
95	85.4	14.2	1602	6	AK024161	Homo sapi
96	60	10.0	60	6	CQ553517	Sequence
97	47.4	7.9	7095	6	AX700181	Sequence
98	47.4	7.9	7095	10	RRMAP1B5	X60370 R.norvegicu
99	47.4	7.9	225204	2	AC127668	Rattus no
100	47.4	7.9	229705	2	AC130178	Rattus no

ALIGNMENTS

RESULT 1
AF070671 1892 bp mRNA linear PRI 21-JUN-1999
DEFINITION Homo sapiens TNF-induced protein G92-1 mRNA, complete cds.
VERSION AF070671.1 GI:3978237
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1892)
ten Cate,J.W., and Pannekoek,H.
Vascular endothelial genes that are responsive to tumor necrosis
factor-alpha in vitro are expressed in atherosclerotic lesions,
including inhibitor of apoptosis protein-1, stannin, and two novel
genes
Blood 93 (10), 3418-3431 (1999)

JOURNAL MEDLINE PUBMED
99252096
10233894
REFERENCE 2 (bases 1 to 1892)
AUTHORS Horrevorts,A.J.G., Fontijn,R.D. and Pannekoek,H.
TITLE Direct Submission
Submitted (05-JUN-1998) Biochemistry, Academic Medical Center,
Meibergdreef 15, Amsterdam 1105 AZ, The Netherlands
location/Qualifiers

FEATURES
source
1..1892
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="endothelial"
/tissue_type="umbilical vein"
98..664
/codon_start=1
/product="TNF-induced protein G92-1"
/protein_id="AAC83229.1"
/db_xref="GI:3978238"
/translation="MATDVNSKNLAOVAOKILKMSKSIATLLIDTSEVDEL
YRVTREYONKKEAKITIKNLITKYLKALILRNQFNDDELALMEKKEKYHQIANT
VVSFHOVDYTFDRNLISRLNECRMELQIILORHLTAKSHGVNVPFHFSDCEFLAA
LYNPGNFKPHQLKLDGINKMLDEBNT"

CDS
ORIGIN
Query Match 97.9%; Score 588.4; DB 9; Length 1892;
Best Local Similarity 99.7%; Pred. No. 2.1e-136;
Matches 600; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTTGACGCTCCGGGCGCCGCTGC-CGACTCTCCGATGGCCACAGATGCTTTAATTC 59
DB 63 CTTGACGCTCCGGGCGCCGCTGC-CGACTCTCCGATGGCCACAGATGCTTTAATTC 122
QY 60 AAAACCTGGCGCTTCAAGCACAAGAGATCTTGGTAAATGTGTCCAATCATCG 119
DB 123 AAAACCTGGCGCTTCAAGCACAAGAGATCTTGGTAAATGTGTCCAATCATCG 182
QY 120 CCACCACTTAATTAACGACACAAGTAGAGTGTGATGATGCTCTTACAGAGTACCA 179
DB 183 CCACCACTTAATTAACGACACAAGTAGAGTGTGATGATGCTCTTACAGAGTACCA 242

QY 180 GGAAGTACACCCAAAACAGAAGAGGAGGAGAGATCATCAAGAACCTCATAGACAG 239
DB 243 GGAATTTACACCCAAAACAGAAGAGGAGGAGAGATCATCAAGAACCTCATAGACAG 302
QY 240 TCATCAAGCTGGCCATCTTTATAGGAATTAATCACTTAATCAAGATGACATTCGA 299
DB 303 TCATCAAGCTGGCCATCTTTATAGGAATTAATCACTTAATCAAGATGACATTCGA 362
QY 300 TGGAGAAATTTAAGAGAAATTCATCAGCTTCGATNACCGTGTGATTTCCATCAGG 359
DB 363 TGGAGAAATTTAAGAGAAATTCATCAGCTTCGATNACCGTGTGATTTCCATCAGG 422
QY 360 TGGATTTATACCTTTGACCGGATGTGTATCCAGCTTTAATGAATGACAGATGAC 419
DB 423 TGGATTTATACCTTTGACCGGATGTGTATCCAGCTTTAATGAATGACAGATGAC 482
QY 420 TGCACCAATATCATTCAGGCGCACTGCTGCAAGTCAATGACGAGGTTAATATGCT 479
DB 483 TGCACCAATATCATTCAGGCGCACTGCTGCAAGTCAATGACGAGGTTAATATGCT 542
QY 480 TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTA 539
DB 543 TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTA 602
QY 540 AACCCACTTACAAAACTATGTGTATGATCAACAAATGTTGATGAGAGACATAT 599
DB 603 AACCCACTTACAAAACTATGTGTATGATCAACAAATGTTGATGAGAGACATAT 662
QY 600 GA 601
DB 663 GA 664

RESULT 2
BD135356
LOCUS
DEFINITION 110 human secretory proteins.
ACCESSION BD135356.1 GI:23230301
VERSION JP 2002508167-A/107.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 1943)
Soppet,D.R., Ruben,S.M., Cater,K.C., Shi,Y., Rosen,C.A.,
Florence,C., Greene,J.M., Feng,P., Ferrie,A.M., Yu,G.L., Janat,F.
and N.J.J.
110 human secretory proteins
Patent: JP 2002508167-A 107 19-MAR-2002;
HUMAN GENOME SCIENCES INC
OS Homo sapiens (human)
PN JP 2002508167-A/107
PD 19-MAR-2002
PR 17-DEC-1998 JP 2000539040
PR 18-DEC-1997 US 60/068 057,18-DEC-1997 US 60/068 007 PR
18-DEC-1997 US 60/068 008,18-DEC-1997 US 60/068 054 PR
18-DEC-1997 US 60/068 064,18-DEC-1997 US 60/068 053 PR
19-DEC-1997 US 60/068 169,19-DEC-1997 US 60/068 368 PR
19-DEC-1997 US 60/068 367,19-DEC-1997 US 60/068 369 PR
19-DEC-1997 US 60/068 365
PI PAUL A MOORE,STEVEN M RUBEN,KENNETH C CARTER,YANGGU SHI,CRAIG
PI A ROSEN,
PI DANIEL R SOPPET,HARA CAOU,YING FEI WEI,KIMBERLY FLORENCE, PI
ROSAANNE D DUAN,
PI CHARLES FLORENCE,JOHN M GREENE,PING FENG,ANN M FERRIE,GUO PI
LIANG YU
PI FORD JANAT,JIAN NI
PI C12N15/09,A61K38/00,A61K48/00,A61P9/00,A61P9/10,A61P15/00, PC
A61P25/00,
PC A61P25/02,A61P25/14,A61P25/16,A61P25/18,A61P25/22,A61P25/24,

	PC	A61P25/28,	
	PC	A61P29/00,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC	
	10,		
	PC	C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53,C12N15/00, PC	
	A61K37/02,		
	PC	C12N5/00	
	CC	n equals a,t,g,o,r c	
	CC	n equals a,t,g,o,r c	
	CC	n equals a,t,g,o,r c	
	CC	n equals a,t,g,o,r c	
	FH	Key	Location/Qualifiers
	FT	source	1..1943
			/organism='Homo sapiens (human)'
FEATURES			
source		Location/Qualifiers	
		1..1943	
		/organism='Homo sapiens'	
		/mol_type='genomic DNA'	
		/db_xref='taxon:9606'	
ORIGIN			
Query Match		97.9%; Score 588.2; DB 6; Length 1943;	
Best Local Similarity		99.3%; Pred. No. 2,4e-136;	
Matches	590; Conservative	0; Mismatches	4; Indels 0; Gaps 0;
OY	8	GTCGCCGCGCGCTGCAGCATCTCTCCGATGGCCACAGATGTCTTTAATTCAAAAACTG	67
Dd	93	GTCCCGCGCGTGCAGCATCTCTCCGATGGCCACAGATGTCTTTAATTCAAAAACTG	152
OY	68	GCCCCATTACGCAAAAGAATCTTGCGTAAAATGATGTCCTTAATTCATGCCACC	127
Dd	153	GCCGTTACGCAAAAGAATCTTGCGTAAAATGATGTCCTTAATTCATGCCACC	212
OY	128	TTAATAGACGACA CAAGTAGAGGTGCTGATGAGCTCTACAGATGACCGAGGAATAC	187
Dd	213	TTAATAGACGACA CAAGTAGAGGTGCTGATGAGCTCTACAGATGACCGAGGAATAC	272
OY	188	ACCCAAAACAAGAGGAGCAGAGAGATCATCAAGAACCTATCAAGACAGTCATCAAG	247
Dd	273	ACCCAAAACAAGAGGAGCAGAGAGATCATCAAGAACCTATCAAGACAGTCATCAAG	332
OY	248	CTGGCCATCTTTAATAGAAATAATCAAGTTTAATCAAGTAGCTAATGATGAGAA	307
Dd	333	CTGGCCATCTTTAATAGAAATAATCAAGTTTAATCAAGTAGCTAATGATGAGAA	392
OY	308	TTTAAGAAGAAGTTCATCAGCTTCTATGACCGGATCAGTTTCATCAGTGGATAT	367
Dd	393	TTTAAGAAGAAGTTCATCAGCTTCTATGACCGGATCAGTTTCATCAGTGGATAT	452
OY	368	ACCTTTGACCGGAATGTTTATCCAGGCTGTTAATGAATCAGAGATCTCACCAA	427
Dd	453	ACCTTTGACCGGAATGTTTATCCAGGCTGTTAATGAATCAGAGATCTCACCAA	512
OY	428	ATCATTCAGCGCCACCTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT	487
Dd	513	ATCATTCAGCGCCACCTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT	572
OY	488	TTTTACAGTTGGAATTTTGGCTGCTGATGATATCTTTGGGAATTTTAAACCCAC	547
Dd	573	TTTTACAGTTGGAATTTTGGCTGCTGATGATATCTTTGGGAATTTTAAACCCAC	632
OY	548	TTACAAAACATATGATGATATCAACAAAATGTTGATGAGAAGACATATGA	601
Dd	633	TTACAAAACATATGATGATATCAACAAAATGTTGATGAGAAGACATATGA	666
RESULT 3			
BD156880			
LOCUS	BD156880	1921 bp	DNA linear PAT 17-JAN-2003
DEFINITION	Primer for synthesizing full-length cdna and use thereof.		
ACCESSION	BD156880.1	GI:27862638	
VERSION	UP 2002191363-A/11723.		
KEYWORDS			

SOURCE ORGANISM	REFERENCE AUTHORS	TITLE JOURNAL	COMMENT
Human sapiens (human)	Okada, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J., Ishii, S., Sugiya, T., Wakamatsu, A., Nagai, K. and Otsuki, T.	Primer for synthesizing full-length cDNA and use thereof Patent: JP 2002191363-A 11723 09-JUL-2002; HELIX RESEARCH INSTITUTE	OS Homo sapiens (human) PN JP 2002191363-A/11723
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	PI TOSHIO OTO, TAKAO ISOGAI, TETSUO NISHIKAWA, KOJI HAYASHI, KAORU SAITO, PI JUNICHI YAMAMOTO, SHIZUKO ISHII, TOMOYASU SUGIYAMA, AI WAKAMATSU, PI KEIICHI NAGAI, TETSUJI OTSUKI	PC C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N5/10, PC C12P21/02, C12Q1/68, C12P21/08, G06F17/30, C12N15/00, C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key	PD 09-JUL-2002 PE 28-JUL-2000 JP 2000280990
FEATURES	Location/Qualifiers	Location/Qualifiers	Location/Qualifiers
source	1. 1921	Location/Qualifiers	1. 1921
ORIGIN	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606"	Location/Qualifiers	Location/Qualifiers
Query Match	97.6%; Score 586.8; DB 6; Length 1921;	Location/Qualifiers	Location/Qualifiers
Best Local Similarity	99.5%; Pred. No. 5.3e-116;	Location/Qualifiers	Location/Qualifiers
Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;		Location/Qualifiers	Location/Qualifiers
1 CTTGACGGCTCCGGCGCGCTGC-CGACTCTCCGATGGCCACAGATGCTTTAATTC	59	1 CTTGACGGCTCCGGCGCGCTGC-CGACTCTCCGATGGCCACAGATGCTTTAATTC	59
105 CTTGACGGCTCCGGCGCGCTGC-CGACTCTCCGATGGCCACAGATGCTTTAATTC	164	105 CTTGACGGCTCCGGCGCGCTGC-CGACTCTCCGATGGCCACAGATGCTTTAATTC	164
60 AAAACCTGGCGCTTCAGGACCAAAAGAGATCTTGGGTAATGTGTCCAAATCCATCG	119	60 AAAACCTGGCGCTTCAGGACCAAAAGAGATCTTGGGTAATGTGTCCAAATCCATCG	119
165 AAAACCTGGCGCTTCAGGACCAAAAGAGATCTTGGGTAATGTGTCCAAATCCATCG	224	165 AAAACCTGGCGCTTCAGGACCAAAAGAGATCTTGGGTAATGTGTCCAAATCCATCG	224
120 CCACCACTTTATAGACGACCAAGTGTGAGTGTCTGATGAGCTCTACAGATGACCA	179	120 CCACCACTTTATAGACGACCAAGTGTGAGTGTCTGATGAGCTCTACAGATGACCA	179
225 CCACCACTTTATAGACGACCAAGTGTGAGTGTCTGATGAGCTCTACAGATGACCA	284	225 CCACCACTTTATAGACGACCAAGTGTGAGTGTCTGATGAGCTCTACAGATGACCA	284
180 GGGAGTACACCCAAACAGAGAGGACAGAGAGATCATCAAGAACCTCTCAAGACAG	239	180 GGGAGTACACCCAAACAGAGAGGACAGAGAGATCATCAAGAACCTCTCAAGACAG	239
285 GGGAGTACACCCAAACAGAGAGGACAGAGAGATCATCAAGAACCTCTCAAGACAG	344	285 GGGAGTACACCCAAACAGAGAGGACAGAGAGATCATCAAGAACCTCTCAAGACAG	344
240 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAATCAAGATGAGCTTACATTTGA	299	240 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAATCAAGATGAGCTTACATTTGA	299
345 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAATCAAGATGAGCTTACATTTGA	404	345 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAATCAAGATGAGCTTACATTTGA	404
300 TGGAGAAATTTAAGAAAGTTATATAGCTGTATGACCGGTGCTGATTTCCATACAG	359	300 TGGAGAAATTTAAGAAAGTTATATAGCTGTATGACCGGTGCTGATTTCCATACAG	359
405 TGGAGAAATTTAAGAAAGTTATATAGCTGTATGACCGGTGCTGATTTCCATACAG	464	405 TGGAGAAATTTAAGAAAGTTATATAGCTGTATGACCGGTGCTGATTTCCATACAG	464
360 TGGATTTATACCTTTGACCGGATGTGTATTCAGAGCTGTTAAATGAATGACAGAGATGC	419	360 TGGATTTATACCTTTGACCGGATGTGTATTCAGAGCTGTTAAATGAATGACAGAGATGC	419
465 TGGATTTATACCTTTGACCGGATGTGTATTCAGAGCTGTTAAATGAATGACAGAGATGC	524	465 TGGATTTATACCTTTGACCGGATGTGTATTCAGAGCTGTTAAATGAATGACAGAGATGC	524
420 TGGACCAATATCTTACAGCGCACCTCACCTGCGAAGTACATGACGGGTTAATATGTCT	479	420 TGGACCAATATCTTACAGCGCACCTCACCTGCGAAGTACATGACGGGTTAATATGTCT	479
525 TGGACCAATATCTTACAGCGCACCTCACCTGCGAAGTACATGACGGGTTAATATGTCT	584	525 TGGACCAATATCTTACAGCGCACCTCACCTGCGAAGTACATGACGGGTTAATATGTCT	584
480 TTGATCATTTTTCAGATTGGAATTTTGGGTGCTGTGTAATCTTTTGGGAAATTTTA	539	480 TTGATCATTTTTCAGATTGGAATTTTGGGTGCTGTGTAATCTTTTGGGAAATTTTA	539
585 TTGATCATTTTTCAGATTGGAATTTTGGGTGCTGTGTAATCTTTTGGGAAATTTTA	644	585 TTGATCATTTTTCAGATTGGAATTTTGGGTGCTGTGTAATCTTTTGGGAAATTTTA	644
540 AACCCCACTTACAAAACTATGTGATGATTCACAAAAATGTTGGATGAAGAGAACATAT	599	540 AACCCCACTTACAAAACTATGTGATGATTCACAAAAATGTTGGATGAAGAGAACATAT	599

|||||
Db 645 AACCCACTTACAAAACCTATGTGATGTGATCAACAAATGTTGATGAAAGACATAT 704
Qy 600 GA 601
Db 705 GA 706
RESULT 4
AX877846 1921 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 12751 from Patent EP1074617.
DEFINITION AX877846
ACCESSION AX877846
VERSION AX877846.1 GI:40032582
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Ota, T., Isegaki, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 12751 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source Location/Qualifiers
1..1921
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
140..706
/note="unnamed protein product"
/codon_start=1
/protein_id="CA889903.1"
/db_xref="GI:40032583"
/translation="MATDVFNSTKMLAVQAKKILGRVNSKSIATTLIDPSSSEVLDEL
/translation="MATDVFNSTKMLAVQAKKILGRVNSKSIATTLIDPSSSEVLDEL
VATREYTONKKEAEKIKILIKTVIKILIVRNNOFNDLALMEKFFKKYHOLAMT
VVSFHOVDYTPDRNYSRLINCEKREMLHOLIQRLHLSKSHGVNVPFHFPSCEFLAA
LYNPNRNPRLHOKLCKDGINKLDEBENT"
ORIGIN
Query Match 97.6%; Score 586.8; DB 6; Length 1921;
Best Local Similarity 99.5%; Pred. No. 5.3e-136;
Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
Qy 1 CTTACGCTCCCGCGCCGCTGC-GCACTCTCCGATGGCCACAGATGCTTTAATCCA 59
Db 105 CTTACGCTCCCGCGCCGCTGC-CGATCTCCGATGGCCACAGATGCTTTAATCCA 164
Qy 60 AAAACCTGGCCGTTGAGGCAAAAAGATCTGGTAAATGGTGCATCAATTCATCG 119
Db 165 AAAACCTGGCCGTTGAGGCAAAAAGATCTGGTAAATGGTGCATCAATTCATCG 224
Qy 120 CCACCACTTAATAGACGACACAAAGTAGAGGTGCTGATGAGCTCTACAGAGTACCA 179
Db 225 CCACCACTTAATAGACGACACAAAGTAGAGGTGCTGATGAGCTCTACAGAGTACCA 284
Qy 180 GGGAGTACACCCCAAAACAAGAGAGGACAGAAAGTCAAGAACTTCATCAAGACAG 239
Db 285 GGGAGTACACCCCAAAACAAGAGAGGACAGAAAGTCAAGAACTTCATCAAGACAG 344
Qy 240 TCATCAAGCTGGCCATCTTTATAGAAATPACAGTTTATCAAGATGAGCTAGCATTA 299
Db 345 TCATCAAGCTGGCCATCTTTATAGAAATPACAGTTTATCAAGATGAGCTAGCATTA 404
Qy 300 TGGAGAAATTTAAGAAAGATTCATCAGCTTCTATGACCGTGTCAAGTTCCATCAG 359
Db 405 TGGAGAAATTTAAGAAAGATTCATCAGCTTCTATGACCGTGTCAAGTTCCATCAG 464
Qy 360 TGGATTATACCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGAGACAGAGATGC 419
Db 465 TGGATTATACCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGAGACAGAGATGC 524

Qy 420 TGCACCAATCATTCAGCGCCACCTCACTGCGCAATGACATGACGGGTAAATATGCT 479
Db 525 TGCACCAATCATTCAGCGCCACCTCACTGCGCAATGACATGACGGGTAAATATGCT 584
Qy 480 TTGATCATTTTTACAGATTGTGAATTTTGGCTGCTGTATATATCTTTGGGAATTTTA 539
Db 585 TTGATCATTTTTACAGATTGTGAATTTTGGCTGCTGTATATATCTTTGGGAATTTTA 644
Qy 540 AACCCACTTACAAAACCTATGTGATGTGATCAACAAATGTTGATGAAAGACATAT 599
Db 645 AACCCACTTACAAAACCTATGTGATGTGATCAACAAATGTTGATGAAAGACATAT 704
Qy 600 GA 601
Db 705 GA 706
RESULT 5
AK001931 1921 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ11069 f1s, clone PLACE1004930, highly similar
DEFINITION to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK001931
VERSION AK001931.1 GI:7023502
KEYWORDS oligo cloning; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Saito, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, Y., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahata, K., Murakami, K., Yasuda, T., Iwyanagi, T., Waga, T.,
Shitatori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Nishimori, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimura, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Houchi, T., Kusano, Y., Kanehori, K., Takahashi, F.,
Hara, H., Tanase, T., Nomura, Y., Togiyasu, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Muesashino, K., Yuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Terashima, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Teraishi, M., Suzaki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Omori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mitsuhashi, S.,
Sachon, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamasaki, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegaki, T., and Sugano, S.
TITLE Complete sequencing and characterization of 21,243 full-length
human cDNAs
JOURNAL Nat. Genet. 36 (1), 40-45 (2004)
PUBMED 14702039
REFERENCE
AUTHORS Isegaki, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y.,
Matsunawa, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
Wakamatsu, A., Nakamura, Y., Negahati, K., Masuho, Y. and Sasaki, N.
TITLE MEDO human cDNA sequencing project
JOURNAL Unpublished
AUTHORS 3 (bases 1 to 1921)
REFERENCE Isegaki, T. and Otsuki, T.
TITLE Direct Submission

JOURNAL Submitted (16-FEB-2000) Takao Isogai, Helix Research Institute,

COMMENT

NEBO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES	Location/Qualifiers
Source	1. .1921

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:.9606"
/c1one="PLAC1.004930"
/r1ssue_type="placenta"
/c1one_id="PLAC1"
/note="Cloning vector: pME185FL3"

```

ORIGIN

Query Match	97.6%;	Score 586.8;	DB 9;	Length 1921;
Best Local Similarity	99.5%;	Pred. No. 5.3e-136;		
Matches 599;	Conservative 0;	Mismatches 2;	Indels 1;	Gaps 1;

[illegible]

DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION	BD156785	GI:27862543
VERSION	BD156785.1	
KEYWORDS	JP 2002191363-A/11628.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CORRECTION

FEATURES

SOLIC

ORIGIN

Query Match	93.9%	Score 564.4	DB 6	Length 1729
Best Local Similarity	99.8%	Pred. No. 2.2e-110		
Matches 565, Conservative	0	Mismatches 1	Indels 0	Gaps 0

Oy	36	TGGCCACAGATGCTTTAATTTCCAAAAA	CTGGCCGCTTACGGCACAAAAAGATCTTTGG	95
Db	102	TGGCCACAGATGCTTTAATTTCCAAAAA	CTGGCCGCTTACGGCACAAAAAGATCTTTGG	161
Oy	96	GTAATAATGATGCCAAATCCATGGCCAC	CACTTAATAGACGACACAAGTAGTAGGTGC	155
Db	162	GTAATAATGATGCCAAATCCATGGCCAC	CACTTAATAGACGACACAAGTAGTAGGTGC	221
Oy	156	TGGATAGCTCTACAGAGTACCAAGGAGTAC	CCAAAAACAGAGGAGGCACAGAGA	215
Db	222	TGGATAGCTCTACAGAGTACCAAGGAGTAC	CCAAAAACAGAGGAGGCACAGAGA	281
Oy	216	TCATCAAGAACCTCATCAAGACAGTCAAT	CAAGCTGGCCATTCTTTAAGGAATATAGT	275
Db	282	TCATCAAGAACCTCATCAAGACAGTCAAT	CAAGCTGGCCATTCTTTAAGGAATATAGT	341
Oy	276	TTAATCAAGATGAGCTACGATTGATGAGAA	ATTTAAGAGAAAGTTCATCAGCTTGCTA	335
Db	342	TTAATCAAGATGAGCTACGATTGATGAGAA	ATTTAAGAGAAAGTTCATCAGCTTGCTA	401
Oy	336	TGACCGTGAGCAGTTTCCATCAGGTGATA	TATACCTTGACGGAAATGTGTTATCCAGGC	395
Db	402	TGACCGTGAGCAGTTTCCATCAGGTGATA	TATACCTTGACGGAAATGTGTTATCCAGGC	461
Oy	396	TGTTAAATGATGACAGAGATGCTGCACCA	ATTCATTCAAGCGCCACCTCACTGCCAAGT	455
Db	462	TGTTAAATGATGACAGAGATGCTGCACCA	ATTCATTCAAGCGCCACCTCACTGCCAAGT	521
Oy	456	CACATGACCGGTTAATTAATGTCTTTGAT	CATTTTTTCAGATTGGAATTTTTGGCTGCT	515
Db	522	CACATGACCGGTTAATTAATGTGTGTTGAT	CAATTTTTTCAGATTGGAATTTTTGGCTGCT	581
Oy	516	TGTATAATCCTTTTGGGAATTTTAAACCA	CTTACAAAACTATGTGATGTATCAACA	575

```

|||||
Db 582 TGTATATTCCTTTGGGAAATTTAAACCCCACTTACAAAACATATGTATGATATCAACA 641
QY 576 AATGTTGGATGAGAGAAACATATGA 601
Db 642 AATGTTGGATGAGAGAAACATATGA 667

RESULT 7
AX877673 1729 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 12578 from Patent EP1074617.
DEFINITION AX877673
ACCESSION AX877673
VERSION AX877673.1 GI:40032409
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Ota, T., Isegai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
TITLE Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K., and Otsuki, T.
JOURNAL Primers for synthesizing full-length cDNA and their use
Research Association for Biotechnology (JP)
FEATURES
source location/Qualifiers
1..1729
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
71..667
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE89829.1"
/db_xref="GI:40032410"
/translation="MHSEAESEKAVTDVFNSEKNLAVQAKKILGKMSKSIATLLID
DTSSEVDELVREYTRTEONKEAEKIKLILKTVKLAIIKRNNOFNODELALMEKF
KKKHQILAMTVSPHOVDYTFDRNVLISRLNCREMLHOLIQRHLTASHGKVNIVFD
HSDSEFLAALNPPGNPKPHIQKICDGINKLDEBNT"

ORIGIN
Query Match 93.9%; Score 564.4; DB 6; Length 1729;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 36 TGGCCACAGATGCTTTAATTCCTTAAATCTGCGCGCTTACGACACAAAGATTTGG 95
Db 102 TGGCCACAGATGCTTTAATTCCTTAAATCTGCGCGCTTACGACACAAAGATTTGG 161
QY 96 GTPAAATGAGTTCACAAATCCATGCGCACACCTTATAGACGACACAAGTAGAGAGTGC 155
Db 162 GTPAAATGAGTTCACAAATCCATGCGCACACCTTATAGACGACACAAGTAGAGAGTGC 221
QY 156 TGGATGAGCTCTACAGAGTACACAGAGATACACCCAAAACAAGAGAGCAGAGAGA 215
Db 222 TGGATGAGCTCTACAGAGTACACAGAGATACACCCAAAACAAGAGAGCAGAGAGA 281
QY 216 TCAATCAAGAACTCATCAAGACAGTATCAAGCTGGCCATTTCTTATAGGAATATCACT 275
Db 282 TCAATCAAGAACTCATCAAGACAGTATCAAGCTGGCCATTTCTTATAGGAATATCACT 341
QY 276 TTAATCAAGATGAGCTAGGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGGCA 335
Db 342 TTAATCAAGATGAGCTAGGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGGCA 401
QY 336 TGAACCGTGTCAAGTTCATCAAGTGTATATACCTTTGACCGGAATGTTATCCAGGC 395
Db 402 TGAACCGTGTCAAGTTCATCAAGTGTATATACCTTTGACCGGAATGTTATCCAGGC 461
QY 396 TGTAAATGAATGACAGAGAGATGCTGCACCAATCATTTACAGCGCCACTCAGTCCCAAGT 455
Db 462 TGTAAATGAATGACAGAGAGATGCTGCACCAATCATTTACAGCGCCACTCAGTCCCAAGT 521

```

```

QY 456 CACATGACGGGTTATATATGCTTGTGATCAATTTTTCAGATGTGTAATTTTGGCTGCT 515
Db 522 CACATGACGGGTTATATATGCTTGTGATCAATTTTTCAGATGTGTAATTTTGGCTGCT 581
QY 516 TGTATATTCCTTTTGGGAAATTTAAACCCCACTTACAAAACATATGTATGATATCAACA 575
Db 582 TGTATATTCCTTTTGGGAAATTTAAACCCCACTTACAAAACATATGTATGATATCAACA 641

QY 576 AATGTTGGATGAGAGAAACATATGA 601
Db 642 AATGTTGGATGAGAGAAACATATGA 667

RESULT 8
AK001850 1729 bp mRNA linear PRI 30-JAN-2004
LOCUS Homo sapiens cDNA FLJ10988 f16, clone PLACE1001920, highly similar
DEFINITION to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK001850
VERSION AK001850.1 GI:7023373
KEYWORDS oligo capping, f16 (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Ohtsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Nishimori, K., Ishibashi, T., Yamashita, H.,
Mizokawa, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hoshino, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunuma, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Saeki, M., Togaishi, T.,
Oyama, M., Hata, H., Matanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Sato, T., Shirai, Y., Takahashi, Y., Takagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Maehuo, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
JOURNAL
PUBMED 14702039
REFERENCE
AUTHORS Isegai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,
Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H.,
Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takemoto, J.,
Arita, M., Nabeckura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,
Wakamatsu, A., Nakamura, Y., Nagahari, K., Maehuo, Y. and Oshima, A.
NEO human cDNA sequencing project
Unpublished
3 (bases 1 to 1729)
Isegai, T. and Otsuki, T.
Direct Submission
Submitted (16-FEB-2000) Takao Isegai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@helix.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)

```

COMMENT

NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5'- & 3'-end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES

Location/Qualifiers
1..1729

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLAC1001920"
/tissue_type="placenta"
/clone_id="PLAC1"
/note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 93.8%; Score 564.4; DB 9; Length 1729;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGTTCAAGCACAAGAGAGATCTTG 95
102 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGTTCAAGCACAAGAGATCTTG 161
96 GTAAGATGCTGTCACAAATTCATGCGCCACCACTTAATAGACACACAAAGTAGAGGTGC 155
162 GTAAGATGCTGTCACAAATTCATGCGCCACCACTTAATAGACACACAAAGTAGAGGTGC 221
156 TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAGAGAGGACAGAGAGA 215
222 TGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAGAGAGGACAGAGAGA 281
216 TCATCAGAAGCTCATCAGACAGTATCAAGCTGGCCATCTTTATAGGAATATACAT 275
282 TCATCAGAAGCTCATCAGACAGTATCAAGCTGGCCATCTTTATAGGAATATACAT 341
276 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTCATCAGCTGCTCA 335
342 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTCATCAGCTGCTCA 401
336 TGAACCTGCTCAGTTTCATCAGAGTATATACCTTTGACCGGAATGTGTATCCAGGC 395
402 TGAACCTGCTCAGTTTCATCAGAGTATATACCTTTGACCGGAATGTGTATCCAGGC 461
396 TGAATATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCACTGCCAAGT 455
462 TGAATATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCACTGCCAAGT 521
456 CACATGAGACGGGTTAATTAATGCTTTGATCATTTTTCAGATTTTGGCTGCCCT 515
522 CACATGAGACGGGTTAATTAATGCTTTGATCATTTTTCAGATTTTGGCTGCCCT 581
516 TGAATATCTCTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATGATCAACA 575
582 TGAATATCTCTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATGATCAACA 641
576 AATGTTGATGAGAGAAATATGA 601
642 AATGTTGATGAGAGAAATATGA 667

RESULT 9
BC005352
LOCUS BC005352 1814 bp mRNA linear PRI 24-FEB-2004
DEFINITION Homo sapiens tumor necrosis factor, alpha-induced protein 8, mRNA
(cDNA clone MGC:12451 IMAGE:3997650), complete cds.
ACCESSION BC005352
VERSION BC005352.1 GI:13529163
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 1814)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, K.H., Schenken, C.M., Schluter, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mulhally, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Guarnate, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiy, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schultz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2. (bases 1 to 1814)
Straussberg, R.
Direct Submission
Submitted (27-MAR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomes Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNLI)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxli.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

REMARK

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNLI at: <http://image.llnl.gov>
Series: IRRL Plate: 16 Row: D Column: 8
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 7657123.
Location/Qualifiers
1..1814

FEATURES

source

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12451 IMAGE:3997650"
/tissue_type="Bladder, carcinoma"
/clone_id="NIH MGC_53"
/lab_host="PH105"
/note="Vector: pDNR-LTR"
1..1814
/gene="TNFAIP8"
/note="synonym: G92-1, MDC-3.13, SCC-S2"
/db_xref="locusid:25816"
144..740
/gene="TNFAIP8"
/codon_start=1
/product="TNFAIP8 protein"
/protein_id="AAH05352.1"
/db_xref="GI:13529164"
/db_xref="locusid:25816"
/translation="MSEAESEKSVATDVFNKLAIVQAKILGKMSKSIATTLID
DTSSSEVLDELVRVTRBYTONKKEAEKIIKLITVYKAILYRNNOFNDLALMEKPF

CDS

misc_feature
KKKHQLAMTVVSFHQVDYTFDRNVLRLNCREMTHQIIRHLJAKSHGRNVNFD
HFSDEFLAALYNPGRNFKPHLOKLCIDGINKMLDEBNT
180..737
/gene="TMFAP8"
/note="DUF758"; Region: Domain of unknown function
(DUF758). Family of eukaryotic proteins with unknown
function, which are induced by tumour necrosis factor"
/db_xref="CD:pfam05527"

ORIGIN

Query Match 93.9%; Score 564.4; DB 9; Length 1814;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTCAGGACAAAAAGAGATCTTG 95
DB 175 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTCAGGACAAAAAGAGATCTTG 234
QY 96 GTAAAAATGCTGTCGAAATCATGCGACCACTTAATAGACACAGATGAGTGC 155
DB 235 GTAAAAATGCTGTCGAAATCATGCGACCACTTAATAGACACAGATGAGTGC 294
QY 156 TGGATGAGCTCTACAGAGTACCGAGAGTACACCCAAAACAAGAGGACAGAGA 215
DB 295 TGGACGAGCTCTACAGAGTACCGAGAGTACACCCAAAACAAGAGGACAGAGA 354
QY 216 TCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCAGT 275
DB 355 TCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCAGT 414
QY 276 TTAATCAAGATGAGTCAAGATGAGAGAAATTTAAGAGAAAGTTCACTGCTGCTA 335
DB 415 TTAATCAAGATGAGTCAAGATGAGAGAAATTTAAGAGAAAGTTCACTGCTGCTA 474
QY 336 TGAACCTGCTGCTTCCATCAAGTGGATTAATCCTTTGACCGGAATGTTTACAGGC 395
DB 475 TGAACCTGCTGCTTCCATCAAGTGGATTAATCCTTTGACCGGAATGTTTACAGGC 534
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAAAATTCATTGAGCGCACCTCACTGCCAAGT 455
DB 535 TGTAAATGAATGACAGAGATGCTGCACCAAAATTCATTGAGCGCACCTCACTGCCAAGT 594
QY 456 CACATGACGGGTTAATAATGCTTTGATCAATTTTCAATGTTGAATTTTGGCTGCT 515
DB 595 CACATGACGGGTTAATAATGCTTTGATCAATTTTCAATGTTGAATTTTGGCTGCT 654
QY 516 TGTATATCTCTTTGGGAATTTAAACCCCACTTAACAAAACCTATGTGATCAACA 575
DB 655 TGTATATCTCTTTGGGAATTTAAACCCCACTTAACAAAACCTATGTGATCAACA 714
QY 576 AAATGTTGATGAGAGAACATATGA 601
DB 715 AAATGTTGATGAGAGAACATATGA 740

RESULT 10
AF099935 1915 bp mRNA linear PRI 12-NOV-1998
LOCUS Homo sapiens MDC-3.13 isoform 2 mRNA, complete cds.
DEFINITION AF099935
ACCESSION AF099935.1 GI:3860092
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Niefeld,W. and Meyershan,A.F.
TITLES Identification of cellular factors involved in the differentiation
of dendritic cells
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1915)
AUTHORS Niefeld,W. and Meyershan,A.F.

TITLE Direct Submission
JOURNAL Submitted (13-OCT-1998) Max Planck Institute for Molecular
Genetics, Imnestrasse 73, Berlin 14195, Germany
FEATURES
source
1..1915
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="adherent monocytes treated with GM-CSF and
IL-4"

CDS
85..681
/note="MDC-3.13/2"
/codon_start=1
/product="MDC-3.13 isoform 2"
/protein_id="AAC72975.1"
/db_xref="GI:3860093"
/translation="HSEAESKEVATDVPSKILAVQAKKILGKMSKSLATLID
DTSSVLDLKVRYREYQNKKEAKKIIKNLIKTYIKIILYRNQFNODELAEKF
KKKHQLAMTVVSFHQVDYTFDRNVLRLNCREMTHQIIRHLJAKSHGRNVNFD
HFSDEFLAALYNPGRNFKPHLOKLCIDGINKMLDEBNT"

ORIGIN

Query Match 93.9%; Score 564.4; DB 9; Length 1915;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTCAGGACAAAAAGAGATCTTG 95
DB 116 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTCAGGACAAAAAGAGATCTTG 175
QY 96 GTAAAAATGCTGTCGAAATCATGCGACCACTTAATAGACACAGATGAGTGC 155
DB 176 GTAAAAATGCTGTCGAAATCATGCGACCACTTAATAGACACAGATGAGTGC 235
QY 156 TGGATGAGCTCTACAGAGTACCGAGAGTACACCCAAAACAAGAGGACAGAGA 215
DB 236 TGGATGAGCTCTACAGAGTACCGAGAGTACACCCAAAACAAGAGGACAGAGA 295
QY 216 TCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCAGT 275
DB 296 TCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCAGT 355
QY 276 TTAATCAAGATGAGTCAAGATGAGAGAAATTTAAGAGAAAGTTCACTGCTGCTA 335
DB 356 TTAATCAAGATGAGTCAAGATGAGAGAAATTTAAGAGAAAGTTCACTGCTGCTA 415
QY 336 TGAACCTGCTGCTTCCATCAAGTGGATTAATCCTTTGACCGGAATGTTTACAGGC 395
DB 416 TGAACCTGCTGCTTCCATCAAGTGGATTAATCCTTTGACCGGAATGTTTACAGGC 475
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAAAATTCATTGAGCGCACCTCACTGCCAAGT 455
DB 476 TGTAAATGAATGACAGAGATGCTGCACCAAAATTCATTGAGCGCACCTCACTGCCAAGT 535
QY 456 CACATGACGGGTTAATAATGCTTTGATCAATTTTCAATGTTGAATTTTGGCTGCT 515
DB 536 CACATGACGGGTTAATAATGCTTTGATCAATTTTCAATGTTGAATTTTGGCTGCT 595
QY 516 TGTATATCTCTTTGGGAATTTAAACCCCACTTAACAAAACCTATGTGATCAACA 575
DB 596 TGTATATCTCTTTGGGAATTTAAACCCCACTTAACAAAACCTATGTGATCAACA 655
QY 576 AAATGTTGATGAGAGAACATATGA 601
DB 656 AAATGTTGATGAGAGAACATATGA 681

RESULT 11
CQ726075 1988 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 12009 from Patent WO02068579.
DEFINITION CQ726075
ACCESSION CQ726075.1 GI:42288070
VERSION

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN

Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1
Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
Kits, such as nucleic acid arrays, comprising a majority of other humateoxons or transcripites, for detecting expression and other uses thereof
Patent: WO 02068579-A 12009 06-SEP-2002;
PE Corporation (NY) (US)
Location/Qualifiers
1. 1988
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

Query Match 93.9%; Score 564.4; DB 6; Length 1988;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTCAAGGACAAAAGAATCTTG 95
205 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTCAAGGACAAAAGAATCTTG 264
96 GTAAATGCTGTCCTCAATTCATGCGCACCACTTAATAGACACACAGATGAGGTGC 155
265 GTAAATGCTGTCCTCAATTCATGCGCACCACTTAATAGACACACAGATGAGGTGC 324
156 TGGATGAGCTCTACAGAGTACACAGGAGTACACCCAAAACAGAGAGGACAGAGAGA 215
325 TGGATGAGCTCTACAGAGTACACAGGAGTACACCCAAAACAGAGAGGACAGAGAGA 384
216 TCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATTTCTTAAGAAATATCAGT 275
385 TCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATTTCTTAAGAAATATCAGT 444
276 TTAATCAAGATGAGTACAGTATGAGAGAAATTAAGAGAAATTCATCAGCTTGCTA 335
445 TTAATCAAGATGAGTACAGTATGAGAGAAATTAAGAGAAATTCATCAGCTTGCTA 504
336 TGAACCTGCTCAAGTTTCATCAGGTGATTAATCTTTGACCGGAATGTGTTATCCAGGC 395
505 TGAACCTGCTCAAGTTTCATCAGGTGATTAATCTTTGACCGGAATGTGTTATCCAGGC 564
396 TGTAAATGAATGACAGAGATGCTGCAACCAATCAATTCAGCGCCCACTTCACTGCCAAGT 455
565 TGTAAATGAATGACAGAGATGCTGCAACCAATCAATTCAGCGCCCACTTCACTGCCAAGT 624
456 CACATGAGACGGGTAAATATGCTTTGATGATTTTTCAGATTTTGGTGGCT 515
625 CACATGAGACGGGTAAATATGCTTTGATGATTTTTCAGATTTTGGTGGCT 684
516 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTAACAAAACATGATGATGATCAACA 575
685 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTAACAAAACATGATGATGATCAACA 744
576 AAATGTTGATGAAGAACAATATGA 601
745 AAATGTTGATGAAGAACAATATGA 770

RESULT 12
AF099936 2003 bp mRNA linear PRI 12-NOV-1998
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens MDC-3.13 isoform 1 mRNA, complete cds.
AF099936.1 GI:3860094
Homo sapiens (human)
Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1. (bases 1 to 2003)
Nietfeld, W. and Meyerhans, A.F.
Identification of cellular factors involved in the differentiation of dendritic cells
Unpublished
2. (bases 1 to 2003)
Nietfeld, W. and Meyerhans, A.F.
Direct Submision
Submitted (13-OCT-1998) Max Planck Institute for Molecular Genetics, Ihmestrasse 73, Berlin 14195, Germany
Location/Qualifiers
1. 2003
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_type="adherent monocytes treated with GM-CSF and IL-4"
IL-4"
198. 770
/note="MDC-3.13/1"
/codon_start=1
/product="MDC-3.13 isoform 1"
/protein_id="AA072976.1"
/db_xref="GI:3860095"
/translation="MAVATDVNSKNLAVQAKLIGKVSIAITLLIDYSSVLD ELYRVREYDYNKKEAEKLIKNLIKVILALILYRNQPNODELAMEFKKHVOLA MIVVSPHQVDYDFDRVLSRLINCEBCEHJQIQRHLTKSHQRVNVNVDHSDCSFL AALVNPFGFKXHLQKLDGINKMDEENI"

Query Match 93.9%; Score 564.4; DB 9; Length 2003;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTCAAGGACAAAAGAATCTTG 95
205 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTCAAGGACAAAAGAATCTTG 264
96 GTAAATGCTGTCCTCAATTCATGCGCACCACTTAATAGACACACAGATGAGGTGC 155
265 GTAAATGCTGTCCTCAATTCATGCGCACCACTTAATAGACACACAGATGAGGTGC 324
156 TGGATGAGCTCTACAGAGTACACAGGAGTACACCCAAAACAGAGAGGACAGAGAGA 215
325 TGGATGAGCTCTACAGAGTACACAGGAGTACACCCAAAACAGAGAGGACAGAGAGA 384
216 TCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATTTCTTAAGAAATATCAGT 275
385 TCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATTTCTTAAGAAATATCAGT 444
276 TTAATCAAGATGAGTACAGTATGAGAGAAATTAAGAGAAATTCATCAGCTTGCTA 335
445 TTAATCAAGATGAGTACAGTATGAGAGAAATTAAGAGAAATTCATCAGCTTGCTA 504
336 TGAACCTGCTCAAGTTTCATCAGGTGATTAATCTTTGACCGGAATGTGTTATCCAGGC 395
505 TGAACCTGCTCAAGTTTCATCAGGTGATTAATCTTTGACCGGAATGTGTTATCCAGGC 564
396 TGTAAATGAATGACAGAGATGCTGCAACCAATCAATTCAGCGCCCACTTCACTGCCAAGT 455
565 TGTAAATGAATGACAGAGATGCTGCAACCAATCAATTCAGCGCCCACTTCACTGCCAAGT 624
456 CACATGAGACGGGTAAATATGCTTTGATGATTTTTCAGATTTTGGTGGCT 515
625 CACATGAGACGGGTAAATATGCTTTGATGATTTTTCAGATTTTGGTGGCT 684
516 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTAACAAAACATGATGATGATCAACA 575
685 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTAACAAAACATGATGATGATCAACA 744
576 AAATGTTGATGAAGAACAATATGA 601
|||||

AF099936 2003 bp mRNA linear PRI 12-NOV-1998
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens MDC-3.13 isoform 1 mRNA, complete cds.
AF099936.1 GI:3860094
Homo sapiens (human)
Homo sapiens

Db 745 AATGTTGATGAAGAATATGA 770

RESULT 13
BC007014

LOCUS
DEFINITION
BC007014 Homo sapiens tumor necrosis factor, alpha-induced protein 8, mRNA (cDNA clone MGC:12346 IMAGE:3930240), complete cds.

ACCESSION
BC007014

VERSION
BC007014.1 GI:13937825

KEYWORDS
MGC.

SOURCE
Homo sapiens
Homo sapiens (human)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 2081)
Straussberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Hopkins, R.F., Zeeberg, B., Bueltow, K.H., Scheffer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marsina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carrincci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunatane, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hilyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skelske, U., Small, D.E., Scherch, A., Schein, J.E., Jones, S.J. and Marra, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
2 (bases 1 to 2081)

JOURNAL
PUBMED
12477932

REFERENCE
Straussberg, R.
Direct Submission
Submitted (30-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

AUTHORS
JOURNAL

REMARK
NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

FEATURES
source
1. 2081
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12346 IMAGE:3930240"
/issue_type="Brain, primitive neuroectodermal"
/clone_lib="NIH_MGC_56"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"

gene 1..2081
/gene="TNFAIP8"
/note="synonyms: GG2-1, MDC-3.13, SCC-S2"
/db_xref="LOCUSID:25816"
237..833
/gene="TNFAIP8"
/codon_start=1
/product="TNFAIP8 protein"
/protein_id="AAH07014.1"
/db_xref="GI:13937825"
/db_xref="LOCUSID:25816"
/translation="MSEASESKVAVDVENSKILAVQAQKLIKGVSKSIATLLID
DTSEVLDLRYVTEYQNKKEKIKLITKVITKAILIRNNQFQDEIATLMEKF
KKVHQLMTVVSFHVQVYTDPRNVLSLNECREMHLIIOHILTAHSGRVNVFD
HPSDEFLAALYNPRGNKPHLQKLCDSDINMLDEEN"
273..830
/gene="TNFAIP8"
/note="DUF758; Region: Domain of unknown function (DUF758). Family of eukaryotic proteins with unknown function, which are induced by tumour necrosis factor"
/db_xref="GDP:pfam05527"

misc_feature
273..830
/gene="TNFAIP8"
/note="DUF758; Region: Domain of unknown function (DUF758). Family of eukaryotic proteins with unknown function, which are induced by tumour necrosis factor"
/db_xref="GDP:pfam05527"

ORIGIN

Query Match 93.9%; Score 564.4; DB 9; Length 2081;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTTAATTCGAAAACCTGGCCGCTTCAGGCACAAAAGAGATCTTGG 95
|||||
268 TGGCCACAGATGCTTTTAATTCGAAAACCTGGCCGCTTCAGGCACAAAAGAGATCTTGG 327
|||||
96 GTAAATGATGTCGCAATTCATCGCCACCACTTAATGACGACACAAGATGAGTGTC 155
|||||
328 GTAAATGATGTCGCAATTCATCGCCACCACTTAATGACGACACAAGATGAGTGTC 387
|||||
156 TGGATGAGCTTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAG 215
|||||
388 TGGATGAGCTTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAG 447
|||||
216 TCATCAAGACCTCATCAAGACGATCAAGACGATCAAGACGATCAAGACGATCAAGAC 275
|||||
448 TCATCAAGACCTCATCAAGACGATCAAGACGATCAAGACGATCAAGACGATCAAGAC 507
|||||
276 TTAATCAAGATGAGTACGATGATGAGTACGATGAGTACGATGAGTACGATGAGTACG 335
|||||
508 TTAATCAAGATGAGTACGATGATGAGTACGATGAGTACGATGAGTACGATGAGTACG 567
|||||
336 TGACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 395
|||||
568 TGACCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 627
|||||
396 TGTTAATGAATGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 455
|||||
628 TGTTAATGAATGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 687
|||||
456 CACATGAGACGGGTATATATGCTTGTGATCAATTTTTCAGATGATGATGATGATGATG 515
|||||
688 CACATGAGACGGGTATATATGCTTGTGATCAATTTTTCAGATGATGATGATGATGATG 747
|||||
516 TGTATATATCCTTTTGGGAAATTTTAAACCCCACTTCAAAAACCTATGATGATGATCA 575
|||||
748 TGTATATATCCTTTTGGGAAATTTTAAACCCCACTTCAAAAACCTATGATGATGATCA 807
|||||
576 AATGTTGATGAAGAATATATGA 601
|||||
808 AATGTTGATGAAGAATATATGA 833
|||||

RESULT 14
AK097284 2502 bp mRNA linear PRI 30-JAN-2004
LOCUS AK097284
DEFINITION Homo sapiens cDNA FLJ39965 f1s, clone SPLEN2027157, highly similar to Homo sapiens MDC-3.13 isoform 2 mRNA.

ACCESSION AK097284
VERSION AK097284.1 GI:21756982
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahari, K., Murakami, K., Yasuda, T., Iwano, T., Magatsuna, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, S., Omura, Y., Abe, K., Kamihara, K., Katsura, N., Sato, K., Tanikawa, M., Yamazaki, M., Nimomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirao, K., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Hosoda, M., Hottu, T., Kuwano, J., Kanehori, K., Takahashi, F., Hara, H., Tanase, T., Nomura, Y., Togiyasu, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuki, H., Oshima, A., Sasaki, N., Aotsuma, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiya, H., Sato, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kikkawa, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Oaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Otsu, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okunura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuno, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T., and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL 14702039

TITLE 2
JOURNAL Kanehori, K., Ishibashi, T., Chiba, Y., Fujimori, K., Hirao, K., S., Tanai, H., Watanabe, S., Ishida, S., Ota, Y., Hottu, T., Wakamatsu, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Ota, T., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H., O., Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Magatsuna, M., Takahashi, F., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K., and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2502)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FUJ Project (HRI Team); 2-6-7 Kazuwa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan, cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'- end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
Location/Qualifiers
1. 2502
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SPLEN2027157"
/tissue_type="spleen"

ORIGIN /clone 11b="SPLEN2"
/note="Cloning vector: pME18SRL3"

Query Match 93.9%; Score 564.4; DB 9; Length 2502;
Best Local Similarity 99.8%; Pred. No. 2.2e-130;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCACAAACCTGCGCTTCAGGACACAAAGATCTTGG 95
720 TGGCCACAGATGCTTTAATTCACAAACCTGCGCTTCAGGACACAAAGATCTTGG 779
96 GTAAATGCTGCTCAATTCATGCGACCACTTAATGACACAGATGAGGTGC 155
780 GTAAATGCTGCTCAATTCATGCGACCACTTAATGACACAGATGAGGTGC 839
156 TGGATGAGCTCTACAGATGACACAGAGATGACACCAACAAAGAGAGGACAGAGA 215
840 TGGATGAGCTCTACAGATGACACAGAGATGACACCAACAAAGAGAGGACAGAGA 899
216 TCATCAAGAACTTCATCAAGACATCATCAAGCTGCGCATCTTTATGGAATATCATG 275
900 TCATCAAGAACTTCATCAAGACATCATCAAGCTGCGCATCTTTATGGAATATCATG 959
276 TTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 335
960 TTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1019
336 TGACCGTGGTCACTTCATCAAGATGATGATGATGATGATGATGATGATGATGATG 395
1020 TGACCGTGGTCACTTCATCAAGATGATGATGATGATGATGATGATGATGATGATG 1079
396 TGTAAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 455
1080 TGTAAATGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139
456 CACATGAGGCGTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 515
1140 CACATGAGGCGTTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
516 TGTATATCTTTTGGAAATTTAAACCCCACTTAACAAATGATGATGATGATGATG 575
1200 TGTATATCTTTTGGAAATTTAAACCCCACTTAACAAATGATGATGATGATGATG 1259
576 AAATGTTGATGAAGAAATATGATGATGATGATGATGATGATGATGATGATGATG 601
1260 AAATGTTGATGAAGAAATATGATGATGATGATGATGATGATGATGATGATGATG 1285

RESULT 15
AC035144
LOCUS 112626 bp DNA linear PRI 27-APR-2001
DEFINITION Homo sapiens chromosome 5 clone CTD-2220M12, complete sequence.
AC035144
VERSION AC035144.3 GI:13811899
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE DOE Joint Genome Institute
JOURNAL Unpublished
2 (bases 1 to 112626)
DOE Joint Genome Institute.
Direct Submission
Submitted (06-APR-2000) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 112626)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Submitted (27-APR-2001) DOE Joint Genome Institute, 2800 Mitchell

COMMENT

Drive, Walnut Creek, CA 94598, USA
On Apr 27, 2001 this sequence version replaced gi:7712094.
Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu

Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.2.
STS Content:

SHGC-7119 G14211.

FEATURES

source

Location/Qualifiers
1..112626
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2220M12"

ORIGIN

Query Match 93.9%; Score 564.4; DB 9; Length 112626;
Best Local Similarity 99.8%; Pred. No. 2.4e-130;

Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAAATTCGAAACCTGGCCGTTCCAGGCACAAAGAGATCTTG 95

5049 TGGCCACAGATGCTTTAAATTCGAAACCTGGCCGTTCCAGGCACAAAGAGATCTTG 65108

96 GTPAAATGGTCCCAATTCATCGCCACCACTTAATAGACGACAAAGTAGAGTGC 155

65109 GTPAAATGGTCCCAATTCATCGCCACCACTTAATAGACGACAAAGTAGAGTGC 65168

156 TGGATGAGCTCTACAGATGACAGGAGATGACCCCAAAACAAGAGAGAGAGAGA 215

65169 TGGATGAGCTCTACAGATGACAGGAGATGACCCCAAAACAAGAGAGAGAGAGA 65228

216 TCATCAAGAACTTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGGAATATCAGT 275

65229 TCATCAAGAACTTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGGAATATCAGT 65288

276 TTAATCAAGATGAGCTACATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCTA 335

65289 TTAATCAAGATGAGCTACATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCTA 65348

336 TGACCGTGGTCACTTCATCAGATGATGATATATCTTTGACCGGAATGTATCCAGGC 395

65349 TGACCGTGGTCACTTCATCAGATGATGATATATCTTTGACCGGAATGTATCCAGGC 65408

396 TGTAAATGATGACAGAGATGCTGCACCAATTCATTCAGCGCCACCTCAGTCCAGT 455

65409 TGTAAATGATGACAGAGATGCTGCACCAATTCATTCAGCGCCACCTCAGTCCAGT 65468

456 CACATGAGAGGGGTAAATATGCTTTGATGATATTTTCAAGTTGTGAATTTTGGCTGCT 515

65469 CACATGAGAGGGGTAAATATGCTTTGATGATATTTTCAAGTTGTGAATTTTGGCTGCT 65528

516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGATGATGATCAACA 575

65529 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGATGATGATCAACA 65588

576 AAATGTTGATGAGAAACATATGA 601

65589 AAATGTTGATGAGAAACATATGA 65614

RESULT 16

AC026795/c

LOCUS AC026795 156277 bp DNA linear PRI 21-SEP-2001
DEFINITION Homo sapiens chromosome 5 clone CTD-2287122, complete sequence.

AC026795

AC026795 GI:15718543

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission
Unpublished

JOURNAL

2 (bases 1 to 156277)
DOE Joint Genome Institute.

JOURNAL

Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

JOURNAL

3 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission

JOURNAL

Submitted (31-AUG-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

REFERENCE

4 (bases 1 to 156277)
DOE Joint Genome Institute and Stanford Human Genome Center.
Direct Submission

JOURNAL

Submitted (21-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

JOURNAL

On Sep 21, 2001 this sequence version replaced gi:15383783.
Draft Sequence Produced by DOE Joint Genome Institute

COMMENT

www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www-shgc.stanford.edu

Quality: Phrap Quality >=40 99.5% of Sequence;
Estimated Total Number of Errors is 0.4.

FEATURES

source
Location/Qualifiers
1..156277
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2287122"

ORIGIN

Query Match 93.9%; Score 564.4; DB 9; Length 156277;
Best Local Similarity 99.8%; Pred. No. 2.4e-130;

Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAAATTCGAAACCTGGCCGTTCCAGGCACAAAGAGATCTTG 95

128233 TGGCCACAGATGCTTTAAATTCGAAACCTGGCCGTTCCAGGCACAAAGAGATCTTG 128174

96 GTPAAATGGTCCCAATTCATCGCCACCACTTAATAGACGACAAAGTAGAGTGC 155

128173 GTPAAATGGTCCCAATTCATCGCCACCACTTAATAGACGACAAAGTAGAGTGC 128114

156 TGGATGAGCTCTACAGATGACAGGAGATGACCCCAAAACAAGAGAGAGAGAGA 215

128113 TGGATGAGCTCTACAGATGACAGGAGATGACCCCAAAACAAGAGAGAGAGAGA 128054

216 TCATCAAGAACTTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGGAATATCAGT 275

128053 TCATCAAGAACTTCATCAAGACAGTCATCAAGCTGGCCATCTTTATAGGAATATCAGT 127994

276 TTAATCAAGATGAGCTACATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCTA 335

127993 TTAATCAAGATGAGCTACATGATGAGAAATTTAAGAAAGAAATTCATCAGCTTGCTA 127934

336 TGACCGTGGTCACTTCATCAGATGATGATATATCTTTGACCGGAATGTATCCAGGC 395

127933 TGACCGTGGTCACTTCATCAGATGATGATATATCTTTGACCGGAATGTATCCAGGC 127874

396 TGTAAATGATGACAGAGATGCTGCACCAATTCATTCAGGCGCCACCTCAGTCCAGT 455

127873 TGTAAATGATGACAGAGATGCTGCACCAATTCATTCAGGCGCCACCTCAGTCCAGT 127814

456 CACATGAGAGGGGTAAATATGCTTTGATGATATTTTCAAGTTGTGAATTTTGGCTGCT 515

DB 127813 CACATGACGGGTTAATATGTGTCATTTTCAGATTGTAATTTTGGCTGCT 127754
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACCTATGTGATGCTTACAA 575
 DB 127753 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACCTATGTGATGCTTACAA 127694
 QY 576 AAATGTTGATGAAGAGACATATGA 601
 DB 127693 AAATGTTGATGAAGAGACATATGA 127668

RESULT 17
 AC138612/c 158057 bp DNA linear PRI 30-MAY-2003
 LOCUS Homo sapiens BAC clone CTD-2293E22 from 2, complete sequence.
 AC138612
 VERSION AC138612.1 GI:27597054
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS Sulston, J.E. and Wilson, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
 PUBMED 95063792
 MEDLINE 9847074
 REFERENCE
 AUTHORS Tomlinson, C. and Bielicki, L.
 TITLE The sequence of Homo sapiens BAC clone CTD-2293E22
 JOURNAL Unpublished (2001)
 REFERENCE
 AUTHORS Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (11-JAN-2003) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE
 AUTHORS Wilson, R.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAY-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: sapiens@wustl.wustl.edu
 Summary Statistics
 Center project name: H_MS2293E22

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

Clone CTD-2293E22 is from a release of the human BAC library CTD.

The library contains cloned DNA from human sperm. See: Shizuya et al., Proc. Natl. Acad. Sci. USA 89:8794-7 (1992); U-J. Kim et al., Genomics 34:213-8 (1996). The clone is available from Research Genetics, Inc. (<http://www.resgen.com>).
 VECTOR: pBelOBAC11
 Selection: chloramphenicol

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone.

FEATURES

source

1. 158057
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="2"
 /map="2"
 /clone="CTD-2293E22"
 /clone_1b="CTD"
 4. 83
 /rpt_family="MIR"
 300. 719
 /rpt_family="ERV1"
 1321. 1459
 /rpt_family="L1"
 1518. 1547
 /rpt_family="AT_rich"
 1964. 1996
 /rpt_family="L1"
 2532. 2575
 /rpt_family="A_rich"
 2639. 2841
 /rpt_family="L2"
 2859. 2886
 /rpt_family="AT_rich"
 3179. 3201
 /rpt_family="AT_rich"
 4351. 4541
 /rpt_family="MIR"
 5128. 5766
 /rpt_family="L1"
 5767. 6328
 /rpt_family="L1"
 6329. 7499
 /rpt_family="L1"
 7503. 7872
 /rpt_family="L1"
 7920. 8106
 /rpt_family="MALR"
 8621. 8849
 /rpt_family="ERV1"
 9113. 9415
 /rpt_family="Alu"
 12251. 12442
 /rpt_family="L1"
 14323. 14361
 /rpt_family="AT_rich"
 14630. 14899
 /rpt_family="Alu"
 15273. 15541
 /rpt_family="L1"
 15534. 17744
 /rpt_family="L1"
 18238. 18436
 /rpt_family="MER2_type"
 18725. 18973
 /rpt_family="MER2_type"
 19348. 19664
 /rpt_family="L2"
 19666. 19723
 /rpt_family="L2"
 20144. 20192
 /rpt_family="polypurine"
 20419. 20468

```
repeat_region /rpt_family="AT_rich"
20476. .20771
/rpt_family="Alu"
repeat_region 21591. .22250
/rpt_family="L1"
repeat_region 22475. .24483
/rpt_family="L1"
repeat_region 24487. .24515
/rpt_family="AT_rich"
repeat_region 25137. .25196
/rpt_family="(CA)n"
repeat_region 26530. .26604
/rpt_family="L2"
repeat_region 27433. .28325
/rpt_family="L2"
repeat_region 28361. .28469
/rpt_family="MIR"
repeat_region 28603. .28653
/rpt_family="(CATATA)n"
repeat_region 29092. .29113
/rpt_family="AT_rich"
repeat_region 29379. .29680
/rpt_family="Alu"
repeat_region 29871. .30008
/rpt_family="MER1_cype"
misc_feature 30569. .30609
/note="Sequence derived from PCR product of project DNA."
repeat_region 31531. .31588
/rpt_family="GA_rich"
repeat_region 31939. .32107
/rpt_family="ERV1"
repeat_region 32426. .32605
/rpt_family="Alu"
repeat_region 32814. .32878
/rpt_family="MIR"
repeat_region 33458. .33570
/rpt_family="L2"
repeat_region 33887. .33984
/rpt_family="(GAAA)n"
repeat_region 35523. .35595
/rpt_family="AT_rich"
repeat_region 36412. .36602
/rpt_family="MIR"
repeat_region 37311. .37413
/rpt_family="L1"
repeat_region 37414. .37724
/rpt_family="Alu"
repeat_region 37725. .37879
/rpt_family="L1"
repeat_region 38402. .38429
/rpt_family="AT_rich"
repeat_region 38459. .38499
/rpt_family="AT_rich"
repeat_region 39513. .39548
/rpt_family="tRNA-Glu-GAG"
repeat_region 39654. .39929
/rpt_family="L1"
repeat_region 39956. .40148
/rpt_family="MIR"
repeat_region 41023. .41048
/rpt_family="(A)n"
repeat_region 41243. .41431
/rpt_family="L1"
repeat_region 41475. .41561
/rpt_family="L1"
repeat_region 41789. .41811
/rpt_family="AT_rich"
```

Query Match	93.9%;	Score 564.4;	DB 9;	Length 158057;
Best Local Similarity	99.8%;	Pred. No. 2.4e-130;		
Matches 565;	Conservative	0;	Mismatches 1;	Indels 0;
Gaps	0;			

[illegible]


```

Db      342  TTAATTAATGATGCTAGCATGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 401
Qy      336  TGAACCGTGGTCACTTTCCATCAGGTGATTAATCCTTTGACCGAAATGTTATCCAGGC 395
Db      402  TGAACCGTGGTCACTTTCCATCAGGTGATTAATCCTTTGACCGAAATGTTATCCAGGC 461
Qy      396  TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCCACTCAGTCCCAAGT 455
Db      462  TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCCACTCAGTCCCAAGT 521
Qy      456  CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATGTTGATTTTGGCTGCT 515
Db      522  CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATGTTGATTTTGGCTGCT 581
Qy      516  TGTATATCCTTTTGGGAAATTTAAACCCCACTTACAAAATATGATGATGATCAACA 575
Db      582  TGTATATCCTTTTGGGAAATTTAAACCCCACTTACAAAATATGATGATGATCAACA 641
Qy      576  AAT- -GTTGATGATGAGACATATGA 601
Db      642  AATGTTGATGATGAGACATATTA 669

RESULT 20
AX869333
LOCUS      AX869333      816 bp      DNA      linear      PAT 17-DEC-2003
DEFINITION Sequence 4238 from Patent EPI074617.
ACCESSION AX869333
VERSION    AX869333.1
KEYWORDS   GI:40024196
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1
AUTHORS   Ota,T., Isegai,T., Nishikawa,T., Hayaishi,K., Saito,K., Yamamoto,J.,
            Ishi,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
TITLE      Primers for synthesizing full-length cDNA and their use
JOURNAL    Patent: EP 1074617-A 4238 07-FEB-2001;
            Research Association for Biotechnology (JP)
FEATURES   Location/Qualifiers
            1..816
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      91.1%; Score 547.6; DB 6; Length 816;
Best Local Similarity 98.9%; Pred. No. 3.4e-126;
Matches 562; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

Qy      36  TGGCCACAGATGCTTTAATTCAAAACCTGGCGCTTCAGGACAAAAGAGATCTTGG 95
Db      102  TGGCCACAGATGCTTTAATTCAAAACCTGGCGCTTCAGGACAAAAGAGATCTTGG 161
Qy      96  GTTAAATGCTGTCCTCAATTCATGCGCACCTTAATAGACGACAAAGTAGAGAGTGC 155
Db      162  GTTAAATGCTGTCCTCAATTCATGCGCACCTTAATAGACGACAAAGTAGAGAGTGC 221
Qy      156  TGGATAGCTCTACAGAGTACACAGGAGATGACCCAAAACAAAGAGAGAGAGAGT 215
Db      222  TGGATAGCTCTACAGAGTACACAGGAGATGACCCAAAACAAAGAGAGAGAGAGT 281
Qy      216  TCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATCTTTATAGGAATATCAGT 275
Db      282  TCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATCTTTATAGGAATATCAGT 341
Qy      276  TTTAATCAAGTAGCTTAGCTTATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 335
Db      342  TTTAATCAAGTAGCTTAGCTTATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 401
Qy      336  TGACCGTGGTCACTTTCCATCAGGTGATTAATCCTTTGACCGAAATGTTATCCAGGC 395

```

```

Db      402  TGAACCGTGGTCACTTTCCATCAGGTGATTAATCCTTTGACCGAAATGTTATCCAGGC 461
Qy      396  TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCCACTCAGTCCCAAGT 455
Db      462  TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGAGCGCCACTCAGTCCCAAGT 521
Qy      456  CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATGTTGATTTTGGCTGCT 515
Db      522  CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATGTTGATTTTGGCTGCT 581
Qy      516  TGTATATCCTTTTGGGAAATTTAAACCCCACTTACAAAATATGATGATGATCAACA 575
Db      582  TGTATATCCTTTTGGGAAATTTAAACCCCACTTACAAAATATGATGATGATCAACA 641
Qy      576  AAT- -GTTGATGATGAGACATATGA 601
Db      642  AATGTTGATGATGAGACATATTA 669

RESULT 21
AK097884
LOCUS      AK097884      1766 bp      mRNA      linear      PRI 30-JAN-2004
DEFINITION Homo sapiens cDNA FLJ40565 f1s, clone THYMJ2004688, highly similar
            to Homo sapiens MDC-3.13 isoform 2 mRNA.
ACCESSION AK097884
VERSION    AK097884.1
KEYWORDS   GI:21757783
            Oligo cloning, f1s (full insert sequence).
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1
AUTHORS   Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R.,
            Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H.,
            Sekine,M., Oobayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
            Ishi,S., Yamamoto,J., Nihei,T., Shihahara,T., Tanaka,T.,
            Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M.,
            Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T.,
            Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katuta,N., Sato,K.,
            Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H.,
            Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M.,
            Hirakawa,S., Chiba,Y., Iehida,S., Ono,Y., Takiguchi,S., Watanabe,S.,
            Yosida,M., Horuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A.,
            Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R.,
            Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsumawa,H., Ichihara,T.,
            Shihata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S.,
            Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T.,
            Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K.,
            Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M.,
            Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y.,
            Ozaki,K., Hiro,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N.,
            Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T.,
            Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K.,
            Nakajima,Y., Mizuno,T., Morinaga,M., Saeki,M., Togashi,T.,
            Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J.,
            Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K.,
            Nagase,T., Nomura,N., Kikuchi,H., Maeno,Y., Yamashita,R.,
            Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isegai,T. and Sugano,S.
TITLE      Complete sequencing and characterization of 21,243 full-length
            human cDNAs
JOURNAL    Nat. Genet. 36 (1), 40-45 (2004)
PUBMED    14702039
AUTHORS   Oshima,A., Takahashi-Fujii,A., Tanase,T., Imose,N., Takeuchi,K.,
            Arita,M., Muesahino,K., Yuuki,H., Hata,H., Sugiyama,T., Irie,R.,
            Otsuki,T., Sato,H., Ota,T., Wakamatsu,A., Ishi,S., Yamamoto,J.,
            Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K.,
            Yamashita,H., Matsumoto,K., Nakamura,Y., Sekine,M., Kikuchi,H.,
            Kanda,K., Wagatsuma,M., Murakawa,K., Kanehori,K., Sugiyama,A.,

```

Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
 NEDO human cDNA sequencing project
 Unpublished
 3 (bases 1 to 1766)
 Isogai, T. and Yamamoto, J.
 Direct Submission
 Submitted (04-JUL-2002) Takao Isogai, Flj Project (HRI Team); 2-6-7
 Kazuo-Kamatori, Katsuzo, Chiba 292-0612, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of
 Economy, Trade and Industry of Japan; cDNA full insert sequencing:
 Research Association for Biotechnology (RAB); cDNA library
 construction; Helix Research Institute (HRI) (supported by Japan
 Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
 HRI, and Biotechnology Center, National Institute of Technology and
 Evaluation; clone selection for full insert sequencing: HRI and
 RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers
 1..1766
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="THYMD2004688"
 /tissue_type="thymus"
 /clone_id="THYMD2"
 /note="Cloning vector: pME18SFL3"

ORIGIN

Query Match 83.3%; Score 500.8; DB 9; Length 1766;
 Best Local Similarity 93.2%; Pred. No. 1.8e-114;
 Matches 561; Conservative 0; Mismatches 2; Indels 39; Gaps 2;

```

1  CTTGAGGCTCCGCGCGCGCTGC-CGACTCTCCGATGCGCAGCATGCTTTAATTCGA 59
Db  141 CTTGAGGCTCCGCGCGCGCTGC-CGACTCTCCGATGCGCAGCATGCTTTAATTCGA 200
QY  60 AAAAAGCTGCGCTTCAGGCAAAAAGAAATCTTGGTAAATGGTGTCCAAATCCATCG 119
Db  201 AAAAAGCTGCGCTTCAGGCAAAAAGAAATCTTGGTAAATGGTGTCCAAATCCATCG 260
QY  120 CCAGCACTTAATGAGCAGACAGTAGTAGGCTGATGAGTCAAGATGACCA 179
Db  261 CCAGCACTTAATGAGCAGACAGTAGTAGGCTGATGAGTCAAGATGACCA 309
QY  180 GGGAGTAGACCCAAAACAGAGAGAGAGAGATCATCAAGAACCTCATCAAGCAG 239
Db  310 -----CAGAGAAAGTCAATCAAGAACCTCATCAAGCAG 342
QY  240 TCATCAAGCTGCCATCTTTAATGAAATATCATGTTAATCAAGATGACCTAGCATGGA 299
Db  343 TCATCAAGCTGCCATCTTTAATGAAATATCATGTTAATCAAGATGACCTAGCATGGA 402
QY  300 TGGAGAAATTAAGAGAAAGTTCATCAGCTTTCATGACCGGTGTCAAGTTTCCATCAGG 359
Db  403 TGGAGAAATTAAGAGAAAGTTCATCAGCTTTCATGACCGGTGTCAAGTTTCCATCAGG 462
QY  360 TGGATTATACCTTTGACCGGATGTGTATCCAGGCTGTAAATGAATGACAGAGATGC 419
Db  463 TGGATTATACCTTTGACCGGATGTGTATCCAGGCTGTAAATGAATGACAGAGATGC 522
QY  420 TGGACCAATCATTCAGCGCCACCTCATCGCCAGTCAATGACAGCGGTTAATATGTCT 479
Db  523 TGGACCAATCATTCAGCGCCACCTCATCGCCAGTCAATGACAGCGGTTAATATGTCT 582
QY  480 TTGATCATTTTTCAGATTTGGAATTTTGGCTGCTGTATATCCCTTTGGGAATTTTA 539
Db  583 TTGATCATTTTTCAGATTTGGAATTTTGGCTGCTGTATATCCCTTTGGGAATTTTA 642
QY  540 AACCCCACTTCAAAAATCTATGTATGATCAACAAAATGTGTGATGAGAGACATAT 599
Db  643 AACCCCACTTCAAAAATCTATGTATGATCAACAAAATGTGTGATGAGAGACATAT 702

```

QY 600 GA 601
 Db 703 GA 704

RESULT 22
 AC095257
 LOCUS
 DEFINITION
 AC095257
 AC095257.5 GI:24818022
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus
 Rattus norvegicus (Norway rat)
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS
 1 (bases 1 to 237561)
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,
 Allen, C., Allen, H., Alibrooks, S., Amin, A., Anguiano, D.,
 Anyalebech, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
 Biswal, K., Blair, J., Blankenburg, K., Blych, P., Brown, M.,
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,
 Davila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,
 Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,
 Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,
 Gebregorgis, E., Geer, K., Gill, R., Gilly, M., Guerra, M., Guevara, W.,
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,
 Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogue, M.,
 Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A.,
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,
 Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C.,
 Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,
 Liu, J., Liu, M., Liu, Y., London, P., Longacre, S., Lopez, V.,
 Lorenz, L., Louie, H., Lozano, R., Lu, X., Ma, J.,
 Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A.,
 Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E.,
 Mawhinney, S., McLeod, M.P., McNeill, T.Z., Meenan, E.,
 Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S.,
 Morgan, M., Morris, K., Morris, S., Muniasa, M., Murphy, M., Nair, L.,
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwaokelimeh, O., Okwodu, G., Olarinmoye, A., Pal, S., Parks, K.,
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C.,
 Plopper, F., Polinder, A., Popovic, D., Prims, E., Pu, L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Register, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Rigg, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.,
 Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H.,
 Shetty, J., Shvartbeyn, A., Sisson, I., Sitter, C.D., Smajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uman, K.,
 Vales, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczek, R., Woodson, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 JOURNAL
 TITLE
 2 (bases 1 to 237561)
 AUTHOR
 TITLE
 JOURNAL
 Submitted (16-SEP-2001) Human Genome Sequencing Center, Department

REFERENCE
AUTHORS
TITLE
JOURNAL

of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 237561)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (09-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

On Nov 9, 2002 this sequence version replaced gi:23264563.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GDPL

Center clone name: CH230-10A12

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 229749 bases at least Q40

Consensus quality: 231209 bases at least Q30

Consensus quality: 232003 bases at least Q20

Estimated insert size: 232575; sum-of-contigs estimation

Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 1 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 237561: contig of 237561 bp in length.

Location/Qualifiers

source
1. 237561
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-10A12"

misc_feature
1. 1156
/note="wgs_end-extension
clone_end:Sp6"

misc_feature
6006..6586
/note="clone boundary
clone_end:Sp6
site:BCORI
end sequence:BH303075"

misc_feature
76320..79576
/note="wgs_contig"
complement(232422..233289)
/note="clone boundary
clone_end:T7
site:BCORI
end sequence:BH303071"

misc_feature
234103..235440
/note="wgs_end-extension
clone_end:T7"

misc_feature 237561
/note="wgs_end-extension
clone_end:T7"

ORIGIN

Query Match 75.6%; Score 454.2; DB 2; Length 237561;
Best Local Similarity 85.9%; Pred. No. 9,4e-103;
Matches 504; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

15 CCGCGTCCGCACTCCTCCGATGGCCACAGATGCTTTAATTCAAAAAACCCTGGCCGTTTC 74
DB 95025 CCTGTGTGTTTATCTCTAGTGGGTACAGATGCTTTAATTCAAAAAACCCTGGCCGTTTC 95084
75 AGGCACAAAGAAGATCTTGGTAAATAGTGTCCAAATCCATGGCCACCACTTAATAG 134
DB 95085 AGGCACAAAGAAGATCTTGGTAAATAGTGTCCAAATCCATGGCCACCACTTATG 95144
135 AGCAGCAAGTAGTAGGTGCTGTAGTGAAGCTTCAAGAGTGACAGGAGTACACCCAAA 194
DB 95145 ATGACACCGACGCGAGGCTCTGGATGAGCTGTACAGGGTGACCAAGAGTACACCCAGA 95204
195 ACAAGAAGAGGACAGAAAGATCATCAGAACTCATAGCACTCATCAAGCTGGCCA 254
DB 95205 ACAAGAAGAGGACAGAAAGATCATCAGAACTCATAGCAAGCGGTCAAGCTGCGCG 95264
255 TCTTTATAGAAATATCATGTTTATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGA 314
DB 95265 TTCTCCAGAGAAACAATCATGTTCAACCAAGACGACTGGCTTGAATGAGAAAGTTCAAGA 95324
315 AGAAAGTTCATCAGCTTGCTATGACCGTGTGACGTTTCCATCAGGTGATTAATACCTTTG 374
DB 95325 GAAAGGTGACACGAGTGGCATGACGATGAGTGTGAGCTTTCAACGAGTAGAGTACACTTTG 95384
375 ACCGAAATGTGTTATCCAGGCTGTTAATGATGAGAGAGATGCTGCACCAATCATC 434
DB 95385 ACCGCAACGTGTGCTCCAGGCTGCAACGAGTGGCGAGAGCTCTGCAACGATCATTC 95444
435 AGCGCACCTCAGTGCAGCAATGACATGACAGCGGTTAATATGCTTTGATCATTTTTCAG 494
DB 95445 AGCGCACCTTACCGCCAGATGTCACGACGCGTTAATACGCTTTGATCATTTTTCAG 95504
495 ATTTGAATTTTGGCTGCTGCTGTAATATCTTTTGGAAATTTAAACCCCACTTACAA 554
DB 95505 ATTTGATTTTGGCTGCTGCTGTAATATCTTTTGGAAATTTAAACCTCACTTACAGA 95564
555 AACTATGTATGCTATCAACAAATGTTGATGAGAGAAACATATGA 601
DB 95565 AACTTGTGACGCGATCAACAAATGTTGATGAGAGAAACATATGA 95611

RESULT 23
BC009090 1728 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus tumor necrosis factor, alpha-induced protein 8, mRNA
DEFINITION (CDNA clone MGC:11714 IMAGR:396593), complete cds.
ACCESSION BC009090
VERSION BC009090.1 GI:14318585
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1728)
Struhsberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buehler,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,A.F., Casavant,T.L.,
Schetz,T.B., Brownstein,M.J., Usdin,T.B., Toshitaki,S.,
Carrinck,P., Prange,C., Rana,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Rhee, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmeiz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schmeiz, A., Schein, J.E., Jones, S.J. and Marz, M.A. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1728)
Strausberg, R.
Direct Submission
Submitted (05-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2530, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CNA Library Preparation: Life Technologies, Inc.
CNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunsaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsesed, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ULNL at: <http://image.llnl.gov>
Series: IRAC Plate: 17 Row: n Column: 4
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 19527361.
Location/Qualifiers

1. 1728
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="MGC:11714 IMAGE:3965693"
/tissue="Mammary tumor. Metallothionin-TGF alpha model. 10 month old virgin mouse. Taken by biopsy."
/clone_11b="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1. 1728
/gene="Tnfai18"
/note="synonyms: Nded, Gg2-1, Sec-2"
/db_xref="locusid:106869"
/db_xref="MGI:2147191"
70. 666
/gene="Tnfai18"
/codon_start=1
/product="TNF-induced protein"
/protein_id="AAH09090.1"
/db_xref="GI:14318586"
/db_xref="locusid:106869"
/db_xref="MGI:2147191"
/translation="MLSEAEPRVATDVNSKNLAVQAKILIGWVSKATLTLLD
DTSEVLDLRYTKYTONKKEARVYIKLITVYIKLAVTHNNQFNDELAIMKRF
KKYVHOLAMTVVSFHOVEYTPPDVNSRLNLNCRRLHETLQHNLTAKSHGRVNVFD
HSDCPFLAALNPFGRFKPQLQCDGINKMLDEENT"

ORIGIN
Query Match 73.9%; Score 444.4; DB 10; Length 1728;
Best Local Similarity 86.6%; Pred. No. 2.3e-100;
Matches 490; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 36 TGGCCACAGATGCTTTAATCCAAAACCTGGCCGTTCCAGGCAAAAAGAATCTTGG 95
Db 101 TGGCTCAGATGCTTTAATCCAAAACCTGGCCGTTCCAGGCAAAAAGAATCTTGG 160
Qy 96 GTAAATGGTGCCAAATCATGCCACCACTTAATAGACGACAAAGTAGAGGTGC 155
Db 161 GCAAAATGGATCAATCATGCGACCAAGCTGATCGACGACGACGAGGTGC 220
Qy 156 TGGATGAGCTCTACAGATGACGAGGATGACACCCAAAACAAGAGGACGAGAGA 215
Db 221 TAGATGAGCTGTACAGGATGACCAAGAGTACACCAAGAGAGGCGGAGAGG 280
Qy 216 TCATCAAGAACTCATCAAGACGATCATCAAGCTGCCTTTTATAGGAATATCAGT 275
Db 281 TCATCAAGAACTCATCAAGACGATCATCAAGCTGCCTTTTATAGGAATATCAGT 340
Qy 276 TTAATCAAGATGAGCTGATGATGATGAGAAATTTAAGAAAGATTCATCAGTCTTA 335
Db 341 TCATCAAGACGAGCTGCGGCTCATGAGAAATTTAAGAAAGATTCATCAGTCTTA 400
Qy 336 TGACCGTGTGCTGATTCATCAGATGATTAATCTTGAACCGGAATGTTATCCAGG 395
Db 401 TGACGCTGTGCTGATTCATCAGATGATTAATCTTGAACCGGAATGTTATCCAGG 460
Qy 396 TGTAAATGATGATGACAGATGCTGACCAATTCATTCAGGCGCATTCATGCGAAGT 455
Db 461 TGCTGAACGAGTCCGAGAGCTCTACAGAGATTCATTCAGGCGCATTCATGCGAAGT 520
Qy 456 CACATGAGCGGCTTAATGCTTTGATGATTCATTCATTCATTCATTCATTCATTCATTC 515
Db 521 CTCACGAGCGGCTTAATGCTTTGATGATTCATTCATTCATTCATTCATTCATTCATTC 580
Qy 516 TGTATATCTCTTTTGGAAATTTAAACCCACTTACAAAACATGATGATGATGATGATGATG 575
Db 581 TGTATATCTCTTTTGGAAATTTAAACCCACTTACAAAACATGATGATGATGATGATGATG 640
Qy 576 AAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
Db 641 AAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 666

RESULT 24
AC120859 184327 bp DNA linear ROD 05-AUG-2004
LOCUS
DEFINITION
MUS MUSCULUS CHROMOSOME 18, CLONE RP24-200D2, COMPLETE SEQUENCE.
AC120859
VERSION
AC120859.12 GI:50979395
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 184327)
Birren, B., Nussbaum, C. and Lander, E.
Mammalian Genome Project
Unpublished
2 (bases 1 to 184327)
Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, D., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gadyana, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Horton, A., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karacas, A., Kelle, C., Lacroque, K., Lamazares, R., Landers, T., Lehoczeky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menees, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,

TITLE
JOURNAL
REFERENCE
AUTHORS

Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schaner, S., Schuppach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trifilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (22-JUN-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

TITLE
JOURNAL
COMMENT

Submitted (05-AUG-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 5, 2004 this sequence version replaced gti:49035085.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@road.mit.edu
 Project Information
 Center project name: L21053
 Center clone name: 200_P_2

FEATURES

source

Location/Qualifiers
 1. 184327
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="18"
 /map="18"
 /clone="RP24-200D2"
 /clone_1ib="RPCT-24 Male Mouse BAC"

misc_feature

1. 4
 /note="clone_boundary
 clone_end:SP6
 site:MboI"

repeat_region

675..700
 /rpt_family="AT_rich"

repeat_region

984..1007
 /rpt_family="AT_rich"

repeat_region

3199..3484
 /rpt_family="B4A"

repeat_region

3506..3534
 /rpt_family="CA)n"

repeat_region

complement(4228..4578)
 /rpt_family="MTE"

repeat_region

4848..5049
 /rpt_family="B3"

repeat_region

5526..5572
 /rpt_family="AT_rich"

repeat_region

6521..6815
 /rpt_family="Lx9"

repeat_region

6890..6921
 /rpt_family="AT_rich"

repeat_region

8693..8735
 /rpt_family="TTCC)n"

repeat_region

8699..8700
 /note="probably CC"

repeat_region

9208..9234
 /rpt_family="TTTG)n"

repeat_region

complement(9236..9335)
 /rpt_family="PB1D7"

repeat_region

complement(9603..9751)
 /rpt_family="B1F"

repeat_region

10571..10599
 /rpt_family="AT_rich"

repeat_region

10637..10660
 /rpt_family="AT_rich"

repeat_region

10951..11146
 /rpt_family="B2_Mm1"

repeat_region

12671..12730
 /rpt_family="TG)n"

repeat_region

13141..13214
 /rpt_family="ID3"

repeat_region

13228..13256
 /rpt_family="PB1D10"

repeat_region

13257..13323
 /rpt_family="ID_B1"

repeat_region

13331..13358
 /rpt_family="AT_rich"

repeat_region

13483..13667
 /rpt_family="B3"

repeat_region

13769..14378
 /rpt_family="LIM4"

repeat_region

14519..14619
 /rpt_family="L1_MM"

repeat_region

complement(14772..15069)
 /rpt_family="L1_MM"

repeat_region

15389..15420
 /rpt_family="AT_rich"

repeat_region

17608..17670
 /rpt_family="TATATG)n"

repeat_region

17773..17801
 /rpt_family="AT_rich"

repeat_region

17825..17850
 /rpt_family="AT_rich"

```

repeat_region      18110..18163
                    /rpt_family="AT_rich"
repeat_region      20210..20731
                    /rpt_family="L1_MM"
repeat_region      20759..20876
                    /rpt_family="RSINE1"
repeat_region      21592..21799
                    /rpt_family="B1_MM"
repeat_region      21818..21837
                    /rpt_family="(A)n"
repeat_region      21916..21949
                    /rpt_family="AT_rich"
repeat_region      22545..22831
                    /rpt_family="Lx7"
repeat_region      22892..23011
                    /rpt_family="B1_MM"
repeat_region      22949..22974
                    /note="single clone coverage"
repeat_region      23095..23294
                    /rpt_family="Lx7"
repeat_region      23912..24049
                    /rpt_family="B1_MM"
repeat_region      24017..24236
                    /rpt_family="WTC"
repeat_region      24238..24355
                    /rpt_family="WTC"

Query Match      73.9%; Score 444.4; DB 10; Length 184327;
Best Local Similarity 85.9%; Pred. No. 2.6e-100;
Matches 493; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

28  TCTCCCATGCGCCACAGATGCTTTATTCACAAAACCTGGCGGTTCAGGCACAAAAGAA 87
Db  144846  TTCTGCGATGCTACAGATGCTTCAATTCACAAAACCTGGCGGTTCAGGCACAAAAGAA 144905

88  GATCTTGCGTAAATGCTGTCAAAATCCATCGCCACCACTTAATAGACACAAGTAG 147
Db  144906  GATCTTGCGCAAAATGCTGTATTCAAATTCATCGCCACCACTGATGACACACACAGAG 144965

148  TGAAGTCTGAGTGAAGCTTACAGAGTACCAAGGAGTACACCCAAAACAAGAGAGG 207
Db  144966  CGAGTCTGAGTGAAGCTGTACAGAGTACCAAGGAGTACACCCAAAACAAGAGAGG 145025

208  AGAAGATCATCAAGAACCTCATCAAGACATCATCAAGTGCCTTTTATAGAA 267
Db  145026  GGAAGAGTCAAGAACCTCATCAAGACATCATCAAGTGCCTTTTATAGAA 145085

268  TAATCAGTTTAATCAAGATGAGTACATGATGAGAAATTAAGAAAGATTCATCA 327
Db  145086  CAATCAGTTTAATCAAGATGAGTACATGATGAGAAATTCAGAAAGATTCATCA 145145

328  GCTTGCTATGACCGTGTGCTAGTTTCCATCAGTGTGATTAATCCTTTCAGCGGAATGTT 387
Db  145146  GCTTGCTATGACCGTGTGCTAGTTTCCATCAGTGTGATTAATCCTTTCAGCGGAATGTT 145205

388  ATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCACTCAC 447
Db  145206  GTCAGGCTGTGTAACAGAGATGCTGCACCAATCATTCAGGCGCACTTCAC 145265

448  TGCAAGTCAATGACGAGGCTTAATAATGCTTTCATTCATTTTCAGATTTGTAATTTT 507
Db  145266  CGCAAGTCAATGACGAGGCTTAATAATGCTTTCATTCATTTTCAGATTTGTAATTTT 145325

508  GAGTGCCTTGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGTGATG 567
Db  145326  GAGTGCCTTGTATATCTTTTGGGAATTTTAAACCTTACAAAACCTATGTGATG 145385

568  TATCAACAATAATGTGATGAAGAAATATGCA 601
Db  145386  CATCAACAATAATGTGATGAAGAAATATGCA 145419

```

```

LOCUS      AJ720906      2009 bp      mRNA      linear      VRT 30-SEP-2004
DEFINITION Gallus gallus mRNA for hypothetical protein, clone 29h8.
ACCESSION  AJ720906
VERSION    AJ720906.1  GI:53136471
KEYWORDS   ORF1.
SOURCE     Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.

REFERENCE
AUTHORS    Caldwell, R.B., Kierzek, A.M., Arakawa, H., Bezubov, Y., Zaim, J.,
            Friedler, P., Kuter, S., Biagodarski, A., Kostowska, D., Koter, M.,
            Plachy, O., Carninci, P., Hayashizaki, Y. and Buerstedde, J.M.
TITLE      Full-length cDNAs from bursa lymphocytes to facilitate gene
            function analysis
JOURNAL     Unpublished
            2 (bases 1 to 2009)
AUTHORS    Caldwell, R.B.
TITLE      Direct Submission
JOURNAL     Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
            Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
            1, D-85764 Neuherberg, GERMANY
            Location/Qualifiers
FEATURES
source
1..2009
    /organism="Gallus gallus"
    /mol_type="mRNA"
    /strain="CB"
    /db_xref="taxon:9031"
    /clone="29h8"
    /cell_type="lymphocyte"
    /issue_type="bursa"
    /clone_lib="rikeni"
    /dev_stage="2 weeks old"
    1..2009
    /locus_tag="RCJMB04_29h8"
    <1..188
    /locus_tag="RCJMB04_29h8"
    189..755
    /locus_tag="RCJMB04_29h8"
    /note="ORF1"
    /codon_start=1
    /product="hypothetical protein"
    /protein_id="CA32565.1"
    /db_xref="GI:53136472"
    /translation="MATDVFNKSLAIQAKKILGKMSKSIATTLIDPSSDVLDEL
    YVETKEYNKKBAEKIKNLIKYVAKLILYRNQFNQDEIALMEKPKKVKQLAKT
    VVSHQVDYTPRPRLPSKLNDCRELIHQIRHLTKSHGRVNVVFDHPSDFPLAA
    LNPFPYKHLQKLCDSGVNRMIDBSNT"
    756..>2009
    /locus_tag="RCJMB04_29h8"

3' UTR

ORIGIN
Query Match      68.0%; Score 408.6; DB 5; Length 2009;
Best Local Similarity 82.5%; Pred. No. 2.1e-91;
Matches 468; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

33  CGATGGCAAGATGCTTTAATTCAAAACCTGGCCGTTGAGGCACAAAAGAAATCT 92
Db  187  CAATGGCAAGATGCTTTCAATTCAAAAGCTTGGCCATTCAGGCCCAAGAAATCTC 246

93  TGGGTAAATGCTGTCCAATCCATCGCCACCACTTAATAGACACAGAATAGTAG 152
Db  247  TTGGGAAATGCTGTCCAATCCATCGCCACCACTTGAATAGATAGACAGATAGT 306

153  TGGTGAATGAGCTCTACAGATGACCAAGGAGTACCCAAAACAAGAGAGCGAGAA 212
Db  307  TTTTAATAGAGCTCTACAGATGACCAAGGAGTACCGAAAACAAGAGAGCGAGAA 366

213  AGATCATCAAGAACTCATCAAGACAGTCAAGTGCATCTTTTATAGAAATATTC 272
Db  367  AGATCATTAAGAACTCATTAATATGCTCTCAAAATGGCAATTCCTACAGAAATATC 426

```

RESULT 25
AJ720906

Oy	273	GGTTAAACAGATGAGCTACATGTATGAGAAATTTAAGAAAGAACTTCATCAGCTTG	332
Db	427	AATTTAAACAAAGATGAAATTCCTCAGTGAAGAAAGTTAAGAAAGAAATTCATCAGCTGG	486
Oy	333	CTATGACCGTGTCAGTTTCATCAGAGTGAATTAATCCTTGACCGGAATGTTATCCA	392
Db	487	CGAAGACGGTGTCACTTTCCATCAGGTGAATTAATCCTTGACAGGAATTTCTTGTC	546
Oy	393	GGCTGTAAATGAATGCAGAGAGATGTCGACCAATCATTTACGGCCACCTCACTGCCA	452
Db	547	AACGTGTAAATGACGTGAGAGAGCTGCTTCATCAGATCATTCACGGGTCACTTAAGTGGCA	606
Oy	453	AGTCACATGACGGGTATTAATATGTCCTTGATCAATTTTCAGATTTGTAATTTTGGCTG	512
Db	607	AATGCATGACAGTGTCAACATGHTGTTGATCACTTCTGATTTGTAAATTTTGGCTG	666
Oy	513	CCTGTATTAATCCTTTTGGAAATTTAAACCCCACTTCAAAAGACTATGATGTATATCA	572
Db	667	CCTGTATTAATCCCTTTGGACCTTAAACTTCATTTCAAAAACCTTTGTATGTGTCA	726
Oy	573	ACAAATGTGATGATGAGAAACAAT	599
Db	727	ACAGATGCTAGATGAGGGAACAAT	753

RESULT	26
LOCUS	BD149633
DEFINITION	BD149633 587 bp DNA linear PAT 17-JAN-2003
ACCESSION	BD149633
VERSION	BD149633.1 GI:27855391
KEYWORDS	JP 2002191363-A/4476.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 587)
AUTHORS	Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,..
TITLE	Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
JOURNAL	Primer for synthesizing full-length cDNA and use thereof Patent: JP 2002191363-A 4476 09-JUL-2002; HELIX RESEARCH INSTITUTE
COMMENT	OS Homo sapiens (human) PN JP 2002191363-A/4476 PD 09-JUL-2002 PF 28-JUL-2000 JP 2000280990 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU PI SAITO, PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU, PI KEIICHI NAGAI,TETSUJI OTSUKI PC C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/ PC 10, PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC Primer for synthesizing full-length cDNA and use thereof FH Key Location/Qualifiers FT source 1..587 FT location/Qualifiers 1..587 /Organism="Homo sapiens" /Organism="Homo sapiens (human)",.. /mol_type="genomic DNA" /db_xref="taxon:9606"
FEATURES	
SOURCE	
ORIGIN	

	Query Match	Similarity	Score	DB #	Length
Db	105	96.6%	394.6	6	587
			Pred. No. 6.5e-8;		
	Matches 422; Conservative	0;	Mismatches 13;	Indels 2;	Gaps 2;
Oy	1	CTTAGCGCTCCGCGCGCGCTGC-CGACTCTCCGATGAGCCACAGATGCTTTAATTCACA	59		
	105	CTTAGCGCTCCGCGCGCGCTGC-CATCTCCGATGAGCCACAGATGCTTTAATTCACA	164		

QY	60	AAAACTGGCCGTTGAGGCACAAAAAGAAAGATCTTGGGTAAATGGTGTCCAAATTCATCG	113
Db	165	AAAACCTGGCCGTTGAGGCACAAAAAGAAAGATCTTGGGTAAATGGTGTCCAAATTCATCG	224
QY	120	CCACCACCTTAATGAGCAGACAAAGTAGTAGAGTCTGATGAGCTCTACAGAGTGACCA	179
Db	225	CCACCACCTTAATGAGCAGACAAAGTAGTAGAGTCTGATGAGCTCTACAGAGTGACCA	284
QY	180	GGGAGTACACCCAAAACAAGAGGAGCAGAGAAAGATCATCAAGAACTCATCAAGACAG	239
Db	285	GGGAGTACACCCAAAACAAGAGGAGCAGAGAAAGATCATCAAGAACTCATCAAGACAG	344
QY	240	TCATCAAGCTGGCCATTCTTTATAGGAATAATCAAGTTAATCAAGATGAGCTAGCATTTGA	299
Db	345	TCATCAAGCTGGCCATTCTTTATAGGAATAATCAAGTTAATCAAGATGAGCTAGCATTTGA	404
QY	300	TGGAGAAATTTAAGAGAAGATTCATGAGCTTGCTATGACCGTGTCAAGTTCCATCCAG	359
Db	405	TGGAGAAATTTAAGAGAAGATTCATGAGCTTGCTATGACCGTGTCAAGTTCCATCCAG	464
QY	360	TGATTTATACCTTTGACCGGAATGTGTTATCCAGGCTTTAAATGAATGACAG - AGAGATG	418
Db	465	GGGGTTATACCTTTGACCGGAATGTGTTATCCAAAGCTTTAAATGAATGACANAAGAAATG	524
QY	419	CTGCACCAATCATCTTCA 435	
Db	525	CTGCACCAAAACANTTA 541	

LOCUS	AX869571	587 bp	DNA	linear	PAT 17-DEC-2003
DEFINITION	Sequence 4476 from Patent EP1074617.				
ACCESSION	AX869571				
VERSION	AX869571.1 GI:40024434				
KEYWORDS	.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1				
AUTHORS	Ota,T., Iisagai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.				
TITLE	Primeres for synthesizing full-length cDNA and their use				
JOURNAL	Patent: EP 1074617-A 4476 07-FEB-2001; Research Association For Biotechnology (JP)				
FEATURES	location/Qualifiers				
Source	1..587 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"				

	Query Match	Similarity	65.7%	Score 394.6	DB 6	Length 587
	Best Local	Similarity	96.6%	Pred. No. 6.5e-86		
	Matches	422	Conservative	0	Mismatches 13	Indels 2
						Gaps 2
QY	1	CTTGAGCCTCCCGGCGCCGTGC-GC-AGACTCTCCGATGCCACAGATGCTTTAATTC	59			
DB	105	CTTGAGGCTCCCGGCGCCGTGCAGACTCTCCGATGCCACAGATGCTTTAATTC	164			
QY	60	AAAACTGGCGCTTCAGGACAAAAGAAAGATCTTGGGTAAATGCTGTCCAAATTCATCG	119			
DB	165	AAAACTTGCGCTTCAGGACAAAAGAAAGATCTTGGGTAAATGCTGTCCAAATTCATCG	224			
QY	120	CGACCACTTAAATGACGACAAAGTAGTAGTGCTGATGAGCTCTACAGAGTAGACCA	179			
DB	225	CGACCACTTAAATGACGACAAAGTAGTAGTGCTGATGAGCTCTACAGAGTAGACCA	284			
QY	180	GGAGTACACCCAAAACAAAGAGAGCGAGAAAGATCATCAAGACCTCTCAAGACAG	239			
DB	285	GGAGTACACCCAAAACAAAGAGAGCGAGAAAGATCATCAAGACCTCTCAAGACAG	344			

Oy	240	CACTAACCTGGCCATTCTTTATAGAAATTAATCACTTAATCAAGATAGAGATCTGA	299
Db	345	TCATTAAGCTGGCCATTCTTTATAGAAATTAATCACTTAATCAAGATAGAGATCTGA	404
Oy	300	TGAGAAATTTAAGAGAAAGTTCACTCAGCTTGATAGACCGTGTGACTGATTCATCAG	359
Db	405	TGAGAAATTTAAGAGAAAGTTCACTCAGCTTGATAGACCGTGTGACTGATTCATCAG	464
Oy	360	TGATTTAATCCTTTGACCGGAATGTGTATATCCAGGCTGTTAATTAATGACG-AGAGATG	418
Db	465	GGGGTTAATCCTTTGACCGGAATGTGTATATCCAGGCTGTTAATTAATGACGAAAGAAATG	524
Oy	419	CTGCACCAATCATTCGA	435
Db	525	CTGCACCAAAACANTTA	541

RESULT 28

LOCUS	CR7606636	1559 bp	mRNA	linear	VRT 15-SEP-2004
DEFINITION	Xenopus tropicalis finished cDNA, clone TEGS053008.				
ACCESSION	CR7606636				
VERSION	CR7606636.1	GI:51966545			
KEYWORDS					
SOURCE	Xenopus tropicalis (Silurana tropicalis)				
ORGANISM	Xenopus tropicalis				

REFERENCE

AUTHORS	TITLE	COMMENT
Amaya, B., Ashmead, J. L., Bonfield, J. K., Croning, M. D. R., Davies, R. M., Francis, M. D., Garrett, N., Gilchrist, M. J., Graffam, D. V., McInerney, S. R., Papadopolulu, N., Rogers, J., Smith, J. C., Taylor, R. G., McInerney, S. R. and Zorn, A. M.	Direct Submission	
Submitted (15-SEP-2004)	Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: trp@sanger.ac.uk	
	Sanger Xenopus tropicalis EST/cDNA project.	

Host: Escherichia coli XL1-blue.

FEATURES

```

/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="TEGG053008"
/clone_lib="XGC-egg"
/dev_stage="egg"

```

ORIGIN

Query Match	59.2%	Score 355.6;	DB 5,	Length 1559;
Best Local Similarity	76.1%;	Pred. No.3.9e-78;		
Matches 452; Conservative	0;	Mismatches 139;	Indels 3;	Gaps 1;

Oy	11	CCGGGCGCGTGGCGGACATCTTCGATGGCCACAGATGTCTTAAATCCAAAACCTGGCC	70
Db	131	CCCGGTCGAGGCGCAGTCAGGCAATGGCAACAGATGTCTTCAGTCCAAAATCTGGCT	190
Oy	71	GTTCAAGCACAAGAGATCTTGGGTAAATGG--TGTCAAATCATGTGCCACACC	127
Db	191	GTACAGGCCCAAAAGAAATCTCTTGGCAAGATGGCGTCATCAAAGTACATAGCAACATCC	250
Oy	128	TTAATANGCACAAGAATGTAGGTGCTGGATGAGCTCTACAGATGACCAAGGAGATAC	187
Db	251	CTTATTTGCGATACCAAGGCGTGAATTTTAAATGAATCTGTACCAATTGACAGAGGAGATAC	310
Oy	188	ACCCAAAACAAGAGGAGCAGAGAAGATCATCAAGAACTCATCAAGACAGTATCAAG	247
Db	311	ACCCAGATGAAGAGAAATCCGAAAAGTTATAAAGAACTTCATCAAAAACGGTAGTACAG	370

Oy	248	TTGGCAGTCTTTTATAGAAATTAATCAAGTTAATCAAAATGAGTCAGATTTGATGAGAA	307
Db	371	CTGGCAGTCTCTTACAGAAATATCAAGTTTATATGAGAAAGATTTGCACTCATGAGAG	430
Oy	308	TTTACAGAGAAAGTTTCATCAGCTGTCTATGACCGTCGTCAAGTTCCATCAGGTGAATTAT	367
Db	431	TTCAAGAGAAAGGTCATCAGCTGGCCATGACTGTGTCAGCTTTTATCAGGTGGATAT	490
Oy	368	ACCTTTGACCGGAATGTGTATTCAGGCTGTTAAATGAATGACAGAGATGCTGCACCAA	427
Db	491	ACCTTTGACCGAAACGTGCTTTCCAAATTTGTTAATGATGTCGCGGAGCTGCTCCATCAG	550
Oy	428	ATCATTTACGGGCCACTTCACCTGCGCAAGTCACATGACGGGTTAATTAATGTCTTGATCAT	487
Db	551	GTCAATACGGCCCATTCACCGCCAAAGTCTCACCGAGGTGTAAATAACGTGTGACCAAC	610
Oy	488	TTTTCAGATTGTGAATTTTGGCGGCGCTGTGATATCCTTTGGGAAATTTTAAACCCGAC	547
Db	611	TTTTCAAATTTGTGAATTTTGGCGGCTGTGACATCCCTTTGAGCCTTTCAAAACAGCAC	670
Oy	548	TTACAAAAAATATGTGATGTATCAACAAAATGTGTGATGAAGAGAACATATG	601
Db	671	CTTAGAGACTCTGTGAGAGATCAACAAAATGTGTGACAGAGCAATATTTAA	724

RESULT 29

BC072904	BC072904	1193 bp	mRNA	linear	VRT 03-AUG-2004
LOCUS					
DEFINITION	IMAGE:5074412), complete cds.				
ACCESSION	BC072904				
VERSION	BC072904.1				
KEYWORDS	MGc.				
SOURCE	Xenopus laevis (African clawed frog)				

222

REFERENCE	AUTHORS
1 (bases 1 to 1193)	Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W

TITLE

Journal of
Dev Dym 325 (4) 384-391 (2002)
initiative

PUBMED
PREFRONTIER

2. (bases 1 to 1193)	
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wanner, L., Shennan, C.M., Schuler, G.D., Altschuler, S.F., Zeeberg, B., Bietow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marasina, K., Farmer, A.A., Rubin, G.M., Hong, J., Stapleton, M., Soares, M.B., Bonaldi, M.F., Casavant, T.L., Greenleaf, E., Brownstein, M.J., Uedini, T.B., Noshayuk, S., Carmichael, P., Prange, C., Raha, S., Loggellano, N.A., Peters, G.J., Abramson, R.D., Mulhaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.W., Worley, K.C., Hale, S.C., Garcia, A.M., Gay, L.J., Huliyil, S.W., Vallalath, D.K., Muzny, D.M., Sodergergen, E.J., Lu, X., Gibbs, R.A., Pahay, J., Helton, E., Ketteman, M., Madan, A., Young, A.C., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Boufford, G.C., Blakesley, R.W., Touchman, J.T., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimmwood, J., Schmutz, R., Myers, R.M., Butterfield, Y.S., Krzywinski, J.M., Skalski, U., Smalish, D.E., Schmechel, A., Schein, J.E., Jones, S.J. and Marra, M.A.	
human and mouse cDNA sequences	
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)	
12477932	
3 (bases 1 to 1193)	
Klein, S. and Gerhardt, D.S.	
Direct Submission	
Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene	

REMARK

Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
 NIH-MGC Project
 Contact: XGC help desk
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Dr. Igor David
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca

Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Telia Olson, Diana Palmquist, Anca Petrescu, Anna Lissa Prabhu, Parvaneh Saeedi, JR Santos, Angeliq Scherch, Ursula Skalska, Duane Smalls, Jeff Scott, Miranda Tsai, George Yang, Jacque Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAX Plate: 152 Row: J Column: 4
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, similarity but not identity to protein.

FEATURES

source

1. 1193
 location/Qualifiers

/organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="MGC:80354 IMAGE:5074412"
 /tissue_type="Ovary, Xenopus"
 /clone_lib="NICH_XGC_Ov1"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1. 1193
 /gene="MGC80354"
 /db_xref="LocusID:443963"
 153. 722
 /gene="MGC80354"
 /codon_start=1
 /product="MGC80354 protein"
 /protein_id="AAH72904.1"
 /db_xref="GI:49115086"
 /db_xref="LocusID:443963"
 /translation="MATDIFNSKMLAVQAQKILGMASRYATSLIDTSGEVLDE
 LYQVREYQNKDAEKIKLKIYIKLAVLRNNQFNEEIEGLMEKRFKRVHQLAM
 TVVSFQVEYTFDRNVLTKLNECRELLHQVIRHLTAHSHGRVNVVPHFSNCEFLA
 ALVNPFGPKKHLQRLCNGVNMKLDENI"

gene

CDS

ORIGIN

Query Match 58.9%; Score 354; DB 5; Length 1193;
 Best Local Similarity 75.9%; Pred. No. 9.7e-78;
 Matches 451; Conservative 0; Mismatches 140; Indels 3; Gaps 1;

11 CCGGCCCCGTCGCGACTCTCCGATGGCCACGATGTTTAAATCCAAAACCTGGCC 70
 129 CCGGCCCCGTCGCGACTCTCCGATGGCCACGATGTTTAAATCCAAAACCTGGCC 188
 71 GTTCAGGACCAAAAGAGTCTTGGGTAAATG---TGTCCAAATCCATCGCACACC 127
 189 GTTCAGGACCAAAAGAGTCTTGGGTAAATG---TGTCCAAATCCATCGCACACC 248
 128 TTAATGACGACCAAGTAGTAGTGTCTGATGAGCTTACAGAGTGAACAGAGATAC 187
 249 CTGATGATGATGACAGTAGGCGAAGTTCTAGATGAATCTACAAAGTACTAGGAGTAC 308
 188 ACCCAAAACAAAGAGGCGAGAGATCATCAAGAACTCTATCAAGACATCATGAAG 247
 309 ACCCAAAACAAAGAGGCGAGAGATCATCAAGAACTCTATCAAGACATCATGAAG 368

QY 248 CTGGCATCTCTTATAGGAATATCATGTTTAAATCAAGATGAGTACATGATGAGAAA 307
 DB 369 CTGGCATCTCTTATAGGAATATCATGTTTAAATCAAGATGAGTACATGATGAGAAA 428
 QY 308 TTTAAGAAAAGTTCATCAGCTTCTATGACCGTGTGATGCTTTCATCAGGTGATAT 367
 DB 429 TTTAAGAAAAGTTCATCAGCTTCTATGACCGTGTGATGCTTTCATCAGGTGATAT 488
 QY 368 ACCCTTGACCGGAATGTGTATTCAGAGCGCTGTTAAATGAGATGAGATGTCAGCAA 427
 DB 489 ACCCTTGACCGGAATGTGTATTCAGAGCGCTGTTAAATGAGATGAGATGTCAGCAA 548
 QY 428 ATCATTCAGCGCACCTCATCTCCAGATCAGATGACGCGTAAATATGCTTTATCAT 487
 DB 549 GTCAATGACGCGCATCTCATCGGCAAGTCTCAGGACGCTTAAATACGTTTACACAC 608
 QY 488 TTTTCAGATTTGAAATTTTGGCTGCTGTTGATTAATCCTTTTGGAAATTTAAACCCAC 547
 DB 609 TTTTCAGATTTGAAATTTTGGCTGCTGTTGATTAATCCTTTTGGAAATTTAAACCCAC 668
 QY 548 TTTCAAAAACCTATGATGATGATCAACAAAATCTTGATGATGATGATGATGATGA 601
 DB 669 CTTCAGAGACTCTGTACGAGATCAACAAATGCTGATGATGATGATGATGATGATGA 722

RESULT 30
 AF098933
 LOCUS
 DEFINITION
 Homo sapiens head and neck tumor and metastasis related protein
 mRNA, partial cds.
 AF098933
 AF098933.1 GI:6851132

SOURCE
 ORGANISM
 Homo sapiens (human)

REFERENCE
 AUTHORS
 TITLE
 IDENTIFICATION OF A NOVEL TUMOR NECROSIS FACTOR-ALPHA-INDUCIBLE
 GENE, SCC-82, CONTAINING THE CONSENSUS SEQUENCE OF A DEATH EFFECTOR
 DOMAIN OF FAS-ASSOCIATED DEATH DOMAIN-LIKE INTERLEUKIN-
 1BETA-CONVERTING ENZYME-INHIBITORY PROTEIN
 J. Biol. Chem. 275 (4), 2973-2978 (2000)

JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (14-OCT-1998) Radiation Medicine, Georgetown University,
 3970 Reservoir Road NW, Washington, DC 20007, USA

FEATURES

source

1. 515
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="head and neck squamous carcinoma"
 <1. 304
 /codon_start=2
 /product="head and neck tumor and metastasis related
 protein"

CDS

QY 298 GATGGAATAATTGAAGAAAGTTCATCAGCTTCTATGACCGTGTGATGATGATGATGA 357
 DB 357 GATGGAATAATTGAAGAAAGTTCATCAGCTTCTATGACCGTGTGATGATGATGATGATGA 418

ORIGIN

Query Match 50.6%; Score 304; DB 9; Length 515;
 Best Local Similarity 100.0%; Pred. No. 3.1e-65;
 Matches 304; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db	1	GATGAGAAATTTAAGAGAAAGTTCATCAGCGTTCATGACCGTGTCAGTTTCATCA	60
OY	358	GGTGATTATACCTTTGACCGGAATGTGTTTACAGGCTGTTAAATGAATCAGAGAT	417
Db	61	GGTGATTATACCTTTGACCGGAATGTGTTTCCAGGCTGTTAAATGAATCAGAGAT	120
OY	418	GCTGACCAATCATTGACGGCGCACCTCAGCCGCAAGTCACATGGAACGGTAAATGAT	477
Db	121	GCTGACCAATCATTGACGGCGCACCTCAGCCGCAAGTCACATGGAACGGTAAATGAT	180
OY	478	CTTTGATCATTTTTCAGATTGGAATTTTTGGCTGCTTGTATATATCCCTTTGGGAATTT	537
Db	181	CTTTGATCATTTTTCAGATTGGAATTTTTGGCTGCTTGTATATATCCCTTTGGGAATTT	240
OY	538	TAAACCCACATTACAAAATCTATGTGATGTATCAACAAAATGTTGGATGAAGAACAT	597
Db	241	TAAACCCACATTACAAAATCTATGTGATGTATCAACAAAATGTTGGATGAAGAACAT	300
OY	598	ATGA 601	
Db	301	ATGA 304	
RESULT 31			
LOCUS	AC118938/c		
DEFINITION	Mus musculus clone RP24-174015, WORKING DRAFT SEQUENCE, 8 unordered pieces.		
ACCESSION	AC118938		
VERSION	AC118938.3		
KEYWORDS	HTG, HTGS, PHASE1, HTGS_DRAFT, HTGS_FULLTOP.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 164258) Britten, B., Nusbaum, C. and Lander, E.		
TITLE	Mus musculus, clone RP24-174015		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 164258) Britten, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barne, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Collange, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Gao, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C., McCarty, M., McKean, P., McKernan, K., Meldrum, J., Meneus, L., Milova, T., Miange, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Notbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Strause, N., Subramanian, A., Talana, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Travia, N., Triggillo, J., Vassiliev, H., Vidal, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.-J., Young, G., Zainoun, J., Zemke, L., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-APR-2002) Whitehead Institute/MIT Center for Genome Research, 120 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 164258)		
AUTHORS	Britten, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Arachi, H.M., Barne, N., Baetien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J.S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, Z., Hagopian, D., Hagos, B.,		

```

Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Meneus, L., Mihova, T., Mieng, V., Murphy, T., Naylor,
Nuyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramsamy, U., Raymond, K., Retta, R., Rise, C., Rogov,
Roman, J., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith
Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbbs, M.,
Talamas, J., Teafaye, S., Theodore, J., Topham, K., Travers, B.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu,
Wyman, D., Young, G., Zahoun, J., Zemek, L., Zimmer, A. and Zody, M.
Submitted (24-FEB-2003) Whitehead Institute/MIT Center for Genom
Direct Submissions
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 24, 2003 this sequence version replaced g1:28261544.
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: 174.O.15
Center clone name: 174.O.15
-----
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 162606 bases at least Q40
Consensus quality: 162965 bases at least Q30
Consensus quality: 163183 bases at least Q20
Insert size: 164000; agarose-fp
Insert size: 163558; sum-of-consigs
Quality coverage: 11.0 in Q20 bases; agarose-fp
Quality coverage: 11.0 in Q20 bases; sum-of-consigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
29707: contig of 29707 bp in length
29708
29807: gap of 100 bp
29808
32340: contig of 2533 bp in length
32341
32440: gap of 100 bp
32441
37689: contig of 5249 bp in length
37690
37789: gap of 100 bp
37790
104529: contig of 66740 bp in length
104530
104629: gap of 100 bp
115026: contig of 10397 bp in length
115027
115126: gap of 100 bp
115127
131677: contig of 16551 bp in length
131678
131777: gap of 100 bp
131778
158793: contig of 27016 bp in length
158794
158893: gap of 100 bp
158894
164258: contig of 5365 bp in length.
Location/Qualifiers
1.164258
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-174O15"
/clone.lib="RP24-174O15"
/misc_feature
1.29707
/note="assembly fragment
clone_end:5P6

```

```

misc_feature      29808..32340
                    /note="assembly_fragment"
misc_feature      32441..37689
                    /note="assembly_fragment"
misc_feature      37790..104529
                    /note="assembly_fragment"
misc_feature      104630..115026
                    /note="assembly_fragment"
misc_feature      115127..131677
                    /note="assembly_fragment"
misc_feature      131778..158793
                    /note="assembly_fragment"
misc_feature      158894..164258
                    /note="assembly_fragment"
                    clone_end:T7
                    vector_side:right"

```

ORIGIN

```

Query Match      42.8%; Score 257.4; DB 2; Length 164258;
Best Local Similarity 65.7%; Pred. No. 1.7e-53;
Matches 375; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY      31 TCCGATGCCACAGATGCTTTTAATTCAAAAAAGCTGGCGTTACGGCAAAAGAGAT 90
DB      88007 TACCATAGGTCCTCATGTGTTAGTTCAAGATCTCGCCCTTCAAGCCAGAAAGAT 87948

QY      91 CTGGGTAATAGTGTCCAAATCCATCCGACACCTTAATAGACACAAAGTATGA 150
DB      87947 CCTGACCAAGATAGCCAGCAAACTGTGGCAACATGCTGATGATGACACCGACGTA 87888

QY      151 GGTGCTGATGAGCTCTACAGAGTACCAGGAGTACCCCAAAAGAGAGGACAGA 210
DB      87887 GATCTTGTAGAGCTTACAAATACACAGAAATACACCCCAAAAGAGAGGCCCA 87828

QY      211 GAAGATCATCAAGAACCTCATCAAGACAGTCAAGCTGGCCATTCTTTATAGAAATA 270
DB      87827 CAAGATCATGAAGAGATGCTATCAAGGTGGCAATCAAAATGGCATCTCTACCGAACA 87768

QY      271 TCAGTTTATCAAGATGAGTACATGATGAGAAATTAAGAAAGTTCATCATGCT 330
DB      87767 ACGATGAGTCAAGAGAGAGTATTAATGTGGAAATCCCGAAAGAACTGAACAGAC 87708

QY      331 TCGTATGACCGTGTGATGTTCCATGAGTGAATTAACCTTTCAGCCGAAATGTATTC 390
DB      87707 TCGATGACCATGATGATGCTTTCAGAAATGAGTATCTTGAATGAAATGTCTTC 87648

QY      391 CAGGCTGTTAATGAATGACAGAGATGCTGACCAAAATCATTACGCCCACTCATCTGC 450
DB      87647 TAAGCTTTCGATGAGTGAAGAGCTGATGATGAATGATGATGATGATGATGATG 87588

QY      451 CAAGTCAATGACCGGTTAATTAATGCTTTTTCATTTTTCAGATTGGAATTTTGGC 510
DB      87587 CAGAACCCATGACCGATTAACATGCTTCAACCACTTGTGATGATGGAATTTCTTTC 87528

QY      511 TCGCTGTATTAATCCTTTTGGGATTTTAAACCCCACTTAACAAACTATGATGATGAT 570
DB      87527 CACTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 87468

QY      571 CAACAATAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
DB      87467 CAATAAATGTTAGACGACAGATCCTCTGA 87437

RESULT 32
AC107848      195574 bp      DNA      linear      HTG 22-MAR-2004
LOCUS      Mus musculus chromosome 9 clone RP23-321A19 map 9, *** SEQUENCING
AC107848
ACCESSION      AC107848.5 GI:45598920
VERSION      HTG; HTGS_PHASE1; HTGS_FULFILL; HTGS_ACTIVEPIN.
KEYWORDS      Mus musculus (house mouse)
SOURCE

```

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 195574)

AUTHORS

Birren,B., Nusbaum,C. and Lander,E.

JOURNAL

Mus musculus chromosome 9, clone RP23-321A19

REFERENCE

Unpublished

AUTHORS

2 (bases 1 to 195574)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barrera,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,U., Campolano,A., Chang,J., Chazaro,B., Choquet,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearlano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardina,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M., McKernan,K., Meidrim,J., Menus,L., Mihova,T., Mckenzie,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strause,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE

3 (bases 1 to 195574)

AUTHORS

Birren,B., Nusbaum,C., Lander,E., Aboueleil,A., Allen,N., Anderson,M., Arachchi,H.M., Barrera,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choquet,Y., Collymore,A., Cook,A., Cooke,P., Corum,B., Dearlano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardina,S., Galham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., Maclean,C., Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M., Melirim,J., Menus,L., Mihova,T., Menges,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE

Direct Submission

JOURNAL

Submitted (22-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Mar 22, 2004 this sequence version replaced gi:44886703.

COMMENT

All repeats were identified using RepeatMasker:

COMMENT

http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research

COMMENT

Center code: WIBR

COMMENT

Web site: http://www-seq.wi.mit.edu

COMMENT

Contact: sequence_submissions@genome.wi.mit.edu

COMMENT

Project Information

COMMENT

Center project name: 120615

COMMENT

Center clone name: 321_A_19

COMMENT

NOTE: This is a 'working draft' sequence. It currently

COMMENT

consists of 8 contigs. The true order of the pieces

COMMENT

is not known and their order in this sequence record is


```

/lab host="DH10B"
/notG="Vector: pME18S-FL3"
1..3986
/gene="tnfaip8"
/notes="synonyms: MGC55331, zgc:55331"
/db_xref="LocustID:393303"
196..756
/gene="tnfaip8"
/codon_start=1
/product="tumor necrosis factor, alpha-induced protein 8"
/protein_id="AAH52765.1"
/db_xref="GI:31127058"
/db_xref="LocustID:393303"
/translation="MDSFSTKNALQAKKMSKATKTVANLFDIDTSFVLDLYR
VKKEYTRNRKRAOKITKNIKRYVKGLVLRNQGPNNEELAVERRKKVHTLATAY
SFYQIDFTDRRWMSMLNDCEHLHQAINRHLTAASHARINHVFNHFCDFLATLY
GPSEVYRGHLOKICEGVNKMDEGNL"

ORIGIN
Query Match 42.7%; Score 256.4; DB 5; Length 3986;
Best Local Similarity 66.0%; Pred. No. 2.7e-53;
Matches 371; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 40 CACAGATGCTTTAATTCAGAAAACCTGGCCGTTTCAGGACAGAAAAGAGATCTTGGGTAA 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 195 CATGACTCGTTGAGCAGCAAGAACTAGCCCTGAGGGCCAGAGAAAGCTCATGAGCAA 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 AATGCTGTCGAATTCATGCGCCACCTTAATAGACAGACAAAGTAGAGTGTCTGGA 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 255 GATGGGACCAACAGAGTGCCCAACCTTTATAGACAGACAGAGCGAGGTACTGGA 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 TGAAGCTCTACAGATGACAGGAGTAGACCCCAAAACAAGAGAGGAGAGAGATCAT 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 CGAGCTGTACAGATGACAGGAGTAGACCGGCAACCGCAAGAGGCGCAAGAGATCAT 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 CAGAAGCTTCATCAAGACAGTCAAGCTGGCCATTCTTTATAGAAATATGATTTAA 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 375 CAAAACTCATAGATGAGTGTGTCAGAGTGGCGTCTCTACCGCAACGTCATTTCAA 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 280 TCAAGATGAGCTGATTCATGAGAAATTTAAGAAAGATTCATCAGCTTCATGAC 339
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 435 CAACGAGAGCTGAGCTGATTCAGAGCTTTCTTAAGAAAGTGCATACGCTGCGATGAC 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 340 CGTGTGAGTTTCATCAGGTGATTAATACCTTTCAGCGGAATGTGTTTCCAGGCTGT 399
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 495 AGCTGTAGCTTTCACAGATGACATTCCTTTTATGACGCGGTCAATGATATCTACT 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 400 AATGATGAGAGAGATGCTGACCAATATTCAGCGCCACCTCATCTGCCAAGTCACA 459
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 555 CATGACTGCGGAGACTTCTGACACGAGCATCAATCGGCAATCTAACGCGAATCTCA 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 460 TGAAGCGGTAATTAATGCTTTTATGATCAATTTTTCAGATTTGAAATTTTGGCTCTGTA 519
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 615 CGCCCAATCAACCATGCTTCATCATCTTCGCGAATGATGACTTCTCCGACGCTATA 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 TAATCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGATGATTCACAAAT 579
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 675 CGGACTTCAGAAAGTAGACCGCGGCACTTGACAGAAAGATCTGTGAAGAGTCAACAGAT 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 GTTGATGAGAGAAATATGA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 735 GCTGAGCAAGAGGCAATTTGA 756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 34
BX649252 205949 bp DNA linear VRT 06-FEB-2004
LOCUS zebrafish DNA sequence from clone DKEY-49M19 in linkage group 22,
DEFINITION complete sequence.
ACCESSION BX649252
VERSION BX649252.7 GI:42454788
KEYWORDS HTG.
SOURCE Danio rerio (zebrafish)

```

```

ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 205949)
REFERENCE
AUTHORS
TITLE
JOURNAL
Direct Submission
Submitted (06-FEB-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 6, 2004 this sequence version replaced gi:41392451.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/Celegans/wormpep/Clone-derived
zebrafish pUC subclones occasionally display inconsistency over the
length of mononucleotide A/T runs and conserved TA repeats. Where
this is found the longest good quality representation will be
submitted.
Repeat names beginning 'Dr' were identified by the Recon repeat
discovery system (Zhifeng Bao and Sean Eddy, submitted), and those
beginning 'dr' were identified by Rick Waterman (Stephen Johnson
lab, Mashu). For further information see
http://www.sanger.ac.uk/Projects/D_rerio/fishmask.shtml DKEY-49M19
is from a Zebrafish BAC library
VECTOR: pIndigoBAC-5.

FEATURES
source
1..205949
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone_lib="DKEY-49M19"
/clone_lib="Daniokey"

ORIGIN
Query Match 42.7%; Score 256.4; DB 5; Length 205949;
Best Local Similarity 66.0%; Pred. No. 3e-53;
Matches 371; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 40 CACAGATGCTTTAATTCAGAAAACCTGGCCGTTTCAGGACAGAAAAGAGATCTTGGGTAA 99
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69527 CATGACTCGTTGAGCAGCAAGAACTAGCCCTGAGGGCCAGAGAAAGCTCATGAGCAA 69468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 100 AATGCTGTCGAATTCATGCGCCACCTTAATAGACAGACAAAGTAGAGTGTCTGGA 159
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69467 GATGGGACCAACAGAGTGCCCAACCTTTTATAGACAGACAGAGGAGGTACTGGA 69408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 160 TGAAGCTCTACAGATGACAGGAGTAGACCCCAAAACAAGAGAGGCGCAAGATTCAT 219
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69407 CGAGCTGTACAGATGACCAAGAGTAGACCGGCAACCGCAAGAGAGGCCCAAGATCAT 69348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 220 CAGAAGCTTCATCAAGACAGTCAAGCTGGCCATTCTTTATAGAAATATGATTTAA 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db      69347 CAAAACCTCATCAAGATGGGTGCTCAAGTTGGGGCTCTCTACCGCAACGTCATTCACA 69288
Qy      280 TCAGATGAGCTAGCTAGTATGAGAGAAATTTAGAGAAAGTTCATGACGTTGTATGAC 339
Db      69287 CAACGAGAGAGCTAGCTAGTATGAGAGCGTTTTCGTAAAGAGTGCATACCGTGGAGTAC 69228
Qy      340 CCGTGTCACTTTCATCAAGTGTGATTTATACCTTTGACCGGAATGTTATCCAGGCTGT 399
Db      69227 AGCTTTAGCTTTCACAGATCGACTTCATCTTTGATCGACGCTCATGATTAATCTACT 69168
Qy      400 AATGATGACAGAGATGCTGACCAATCATCTCAGGCGCCACTGCGCAAGTACA 459
Db      69167 CAATGACTGCCGTAGTACTTTCGACACGCGCATCATCGCATTAACGCGGAATCTCA 69108
Qy      460 TGACGGGTTAATATATGCTTTTATGATCAATTTTTCAGATGTCGATTTTGGCTGCTTGA 519
Db      69107 CGCCGGAATCAACCATGCTTCATCATTTTCCCGATTTGATGATCTTCCTCGGAGGCTATA 69048
Qy      520 TAATCTTTTGGGAATTTTAAACCCCTTACAAAACCTATGATGATGATCAACAAAT 579
Db      69047 CGACCTTCAGAGATGATCGCGGCACTTCGAGAGATCTGTAGAGAGTCAACAGAT 68988
Qy      580 GTTGATGAGAGAGAACATATGA 601
Db      68987 GCTGACGAGAGGCAATCTTGA 68966

```

RESULT 35

```

AC012678 188937 bp DNA linear HTG 18-JUL-2000
LOCUS     Homo sapiens chromosome 15 clone RP11-394B5, WORKING DRAFT
DEFINITION
VERSION   AC012678.2 GI:6728953
KEYWORDS  HTG; HTGS PHASE; HTGS_DRAFT.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

```

REFERENCE
AUTHORS   Bruno, D., Conn, L., Dela Rosa, M., Federpiel, N., Foreman, P.,
           Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
           Morehouse, A., J., Oefner, P., Palm, C. J., Ramirez, D., Wilhelmy, J.,
           Yu, S. and Davis, R. W.
           Unpublished
JOURNAL   2 (bases 1 to 188937)
REFERENCE
AUTHORS   Bruno, D., Conn, L., Dela Rosa, M., Federpiel, N., Foreman, P.,
           Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R.,
           Morehouse, A., J., Oefner, P., Palm, C. J., Ramirez, D., Wilhelmy, J.,
           Yu, S. and Davis, R. W.
           Direct Submission

```

```

JOURNAL   Submitted (03-NOV-1999) DNA Sequencing and Technology Center,
           Stanford University, 855 California Avenue, Palo Alto, CA 94304,
           USA
TITLE     On Jan 21, 2000 this sequence version replaced gi:6223629.
JOURNAL   --- Genome Center

```

COMMENT

```

Center: Stanford DNA Sequencing and Technology Development
Center code: SDSTDC
Web site: http://sequence-www.stanford.edu/group/human/
Contact: hum-info@sequence.stanford.edu
Project Information
Center project name: 698
Center clone name: RP11-394B5
----- Summary Statistics
Sequencing Vector: M13mp18; X02513
Chemistry: Dye primer, 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 177081 bases at least Q40
Consensus quality: 185332 bases at least Q30
Consensus quality: 186766 bases at least Q20

```

```

Insert size: 185828; agarose-fp
Insert size: 188487; sum-of-ctnigs
Quality coverage: 5.3 in Q20 bases; agarose-fp
Quality coverage: 5.2 in Q20 bases; sum-of-ctnigs.
NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2170: contig of 2170 bp in length
2171: gap of unknown length
2221: contig of 4998 bp in length
7268: gap of unknown length
7269: gap of unknown length
15922: contig of 8654 bp in length
15923: gap of unknown length
15973: contig of 10897 bp in length
26870: gap of unknown length
26920: contig of 12827 bp in length
39746: gap of unknown length
39796: gap of unknown length
39797: contig of 14873 bp in length
54670: gap of unknown length
54720: contig of 15178 bp in length
69897: gap of unknown length
69898: gap of unknown length
69947: gap of unknown length
93475: contig of 23528 bp in length
93476: gap of unknown length
93526: gap of unknown length
138958: contig of 45433 bp in length
138959: gap of unknown length
189008: gap of unknown length
189337: contig of 49929 bp in length.
Location/Qualifiers
1. 188937
/mol_type="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="15"
/clone="RP11-394B5"
/location="RP11 human BAC library 11"

```

ORIGIN

```

Query Match 42.6%; Score 255.8; DB 2; Length 188937;
Best Local Similarity 65.5%; Pred. No. 4.2e-53;
Matches 374; Conservative 0; Mismatches 197; Indels 0; Gaps 0;
Qy      31 TCCGATGCGCAGATGCTTTAATCCAAAACCTGGCGCTTCAGGCAAAAAGAGAT 90
Db      51681 TACCATAGCTCTGATGTTTATGTTCAAAAGATCTTGCGCTTCAAGCCGAGAAAGAT 51740
Qy      91 CTGGGTAAATGATGTCCAATCCATCGCCACACCTTAATAGACGACACAAGTAGTGA 150
Db      51741 TCTGACCAAAATGACGACGAAACTGTGGCAACATGTTGATGATGACACGACGCGA 51800
Qy      151 GGTGCTGATGACCTTACAGAGTGACCGAGGAGTACACCCAAAACAAGAGAGCGAGA 210
Db      51801 GATCTTGTATGAGCTCTACAAAGTCACAAAAGAGCACACACAACAAGAGAGCGCA 51860
Qy      211 GAGATCATCAAGAACCTCATCAAGACAGTCAACGACGCGCATCTTATAGAAATTA 270
Db      51861 CAAGATCATGAAAGACTTATCAAGGTGGGATCAAAATCGGATCCTTACCGGAACAA 51920
Qy      271 TCAGTTTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 330
Db      51921 CAGTTTAAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 51980
Qy      331 TGTATGACCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 390
Db      51981 CGCATGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 52040
Qy      391 CAGGCTGTTAAATGAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 450
Db      52041 CATCTCTGCAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 52100

```


[illegible]

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1
 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 Kites, such as nucleic acid arrays, comprising a majority of humaxons or transcripts, for detecting expression and other uses thereof

JOURNAL Patent: WO 0206879-A 22590 06-SEP-2002;
 PE Corporation (NY) (US)
 Location/Qualifiers
 1..1248
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 42.3%; Score 254.4; DB 6; Length 1248;
 Best Local Similarity 65.9%; Pred. No. 8.2e-53;
 Matches 369; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 42 CAGATGCTTTAAATCCAAAACCTGGCCGTTTCAGGACAAAAGAGATCTGGGTAAA 101
 |||||
 DB 689 CTGATGTTTTTAAAGTCTTGCGCTTCAAGCCAGAGAAATCTGAGCAAAA 748
 |||||

QY 102 TGGTGTCCAAATCATTGCGCAACAATTAAAGACACAAAGTAGTAGGTGCTGATG 161
 |||||
 DB 749 TAGCCAGCAAAACGTGGCCAAACATGTTGATGATGACACAGACGAGATCTTTGATG 808
 |||||

QY 162 AGCTCTACAGAGTAGACAGGAGTAGACCCAAAACAAGAGAGGACAGAAATATCA 221
 |||||
 DB 809 AGCTCTACAAAGTACCAAGAGACACACAAAGAAAGAGCCACAAAGATCA 868
 |||||

QY 222 AGAACCCTCATCAAGACAGTCAAGCTGGCCATTTCTTATAGAAATATCAAGTTATC 281
 |||||
 DB 869 AAGACTTATCAAGAGTGGCATCAAAATCGGATCTCTACCGGACACCAAGTTAGCC 928
 |||||

QY 282 AAGATGAGTACATGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGTATGACG 341
 |||||
 DB 929 AAGAGAGCTGTTATGTTAGAGAAAGTTCCGAAAGAGTGAACACGACCGCATGACCA 988
 |||||

QY 342 TGGTCACTTCCATCAGGTGATTTATACCTTTGACCGGATGTTATCCAGCTGTAA 401
 |||||
 DB 989 TTGTCAAGCTTCTATGAGAGTGAATACCTTCATAGAAAGTGTCTCTCAATCTCTGC 1048
 |||||

QY 402 ATGAATGACAGAGATGCTGCACCAAAATCAATTCAGGCGCACCTCATCTGCCAAGTCACATG 461
 |||||
 DB 1049 ATGAGTGCAGAGACCTGTGTGATGAACTGTGACGCGGACCTGACGCCAGAGACCAAG 1108
 |||||

QY 462 GACGGGTTAATAATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATA 521
 |||||
 DB 1109 GGGGATCAACCAAGCTTTTAAACACTTTCGCGATGTGAGTTCTCTCAACCTCTATA 1168
 |||||

QY 522 ATCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGTGATGATCAACAAAATGT 581
 |||||
 DB 1169 GTCTGATGAGAGCTGTAGGCCCAACCTCAAGAGGATTTGTGAAGATCAATAGTTGC 1228
 |||||

QY 582 TGGATGAAGAGAACATATGA 601
 |||||
 DB 1229 TAGATGAGAAAGTCTTTAA 1248
 |||||

RESULT 38
 CO841622 1986 bp DNA linear PAT 02-AUG-2004
 LOCUS CO841622
 DEFINITION Sequence 269 from Patent EP1440981.
 ACCESSION CO841622
 VERSION CO841622.1 GI:50893409
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J., Isono, Y., Nagai, K. and Irie, R.
 Full-length human cDNA
 Patent: EP 1440981-A 269 28-JUL-2004;
 Research Association for Biotechnology (JP)
 Location/Qualifiers
 1..1986
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 42.3%; Score 254.4; DB 6; Length 1986;
 Best Local Similarity 65.9%; Pred. No. 8.3e-53;
 Matches 369; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 42 CAGATGCTTTAAATCCAAAACCTGGCCGTTTCAGGACAAAAGAGATCTGGGTAAA 101
 |||||
 DB 150 CTGATGTTTTTAAAGTCTTGCGCTTCAAGCCAGAGAAATCTGAGCAAAA 209
 |||||

QY 102 TGGTGTCCAAATCATTGCGCAACAATTAAAGACACAAAGTAGTAGGTGCTGATG 161
 |||||
 DB 210 TAGCCAGCAAAACGTGGCCAAACATGTTGATGATGACACAGACGAGATCTTTGATG 269
 |||||

QY 162 AGCTCTACAGAGTAGACAGGAGTAGACCCAAAACAAGAGAGGACAGAAATATCA 221
 |||||
 DB 270 AGCTCTACAAAGTACCAAGAGACACACAAAGAAAGAGCCACAAAGATCA 329
 |||||

QY 222 AGAACCCTCATCAAGACAGTCAAGCTGGCCATTTCTTATAGAAATATCAAGTTATC 281
 |||||
 DB 330 AAGACTTATCAAGAGTGGCATCAAAATCGGATCTCTACCGGACACCAAGTTAGCC 389
 |||||

QY 282 AAGATGAGTACATGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGTATGACG 341
 |||||
 DB 390 AAGAGAGCTGTTATGTTAGAGAAAGTTCCGAAAGAGTGAACACGACCGCATGACCA 449
 |||||

QY 342 TGGTCACTTCCATCAGGTGATTTATACCTTTGACCGGATGTTATCCAGCTGTAA 401
 |||||
 DB 450 TTGTCAAGCTTCTATGAGAGTGAATACACTTGTGATGAAACGTCCTCAATCTCTGC 509
 |||||

QY 402 ATGAATGACAGAGATGCTGCACCAAAATCAATTCAGGCGCACCTCATCTGCCAAGTCACATG 461
 |||||
 DB 510 ATGAGTGCAGAGACCTGTGTGATGAACTGTGTGAGGAGGACCTGACGCCAGAGACCAAG 569
 |||||

QY 462 GACGGGTTAATAATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATA 521
 |||||
 DB 570 GGGGATCAACCAAGCTTTTAAACACTTTCGCGATGTGAGTTCTCTCAACCTCTATA 629
 |||||

QY 522 ATCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGTGATGATCAACAAAATGT 581
 |||||
 DB 630 GTCTGATGAGAGCTGTAGGCCCAACCTCAAGAGGATTTGTGAAGATCAATAGTTGC 689
 |||||

QY 582 TGGATGAAGAGAACATATGA 601
 |||||
 DB 690 TAGATGAGAAAGTCTTTAA 709
 |||||

RESULT 39
 AK123281 1986 bp mRNA linear PRI 09-SEP-2003
 LOCUS AK123281
 DEFINITION Homo sapiens cDNA FLJ41287 f1s, clone BRAMY2038484, moderately similar to Homo sapiens MDC-3.13 isoform 2 mRNA.
 ACCESSION AK123281
 VERSION AK123281.1 GI:34528784
 KEYWORDS Oligo capping; f1s (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,

Sugiyama, T., Irie, R., Otsubi, T., Saeo, H., Makamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yameshita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanbori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kakaki, B., Suzuki, Y., Sugano, S., Nagahari, K., Nasuto, Y., Nagai, K. and Isogai, T. Nucleo human cDNA sequencing project

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kametarai, Kisarazu, Chiba 292-0818, Japan
Tel: +81-477-450-1000 Fax: +81-477-450-1001
E-mail: isogai@hri.ac.jp

COMMENT

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center.); 5'- & 3'-end one pass sequencing: RAB. HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
SOU

```

/vecid
/source
1..1986
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRAMY2038484"
/tissue_type="amygdala"
/clone_id="BRAMY2"
/note="Cloning vector: pME185FL3"
95..709
CDS

```

CDS

ORIGIN

Query Match	42.3%	Score 254.4	DB 9	Length 1986
Best Local Similarity	65.9%	Pred. No. 8.3e-53		
Matches 369	Conservative	0	Mismatches 191	Indels 0
			Gaps	0

Oy	42	CAGATGCTTTAATTCGAAAAACCTGGCGCTTCAGGACCAAAAGAAAGATCTTGGGTAAA	101
Db	150	CTGATGTTTTTAGTTCAAAGAGTCTTGGCTTCGAAGCCGAGAAGAGATTCTTGACCAAAA	209
Oy	102	TGCTGTCCAAATCCATTCGCCACCACTTAAATGAGCGACCAAGTAGTAGTGAGTGTGGATG	161
Db	210	TAGCCAGCAAAACTGTGGCCCAACATGTTGATTGATGACCAAGCCAGATCTTTGATG	269
Oy	162	AGCTTCAGAGGTGCACGAGGAGTACACCCAAAACAAGAGGCGAAGAAATCATCA	221
Db	270	AGCTTCACAAAGTCAACAAAGAGCACACACAACAAGAAAGAGGCCACAGATCATGA	329
Oy	222	AGAACCTCATCAAGCAGTCATCAAGCTGGCCATTCTTTATAGAAATATCGTTTAATC	281
Db	330	AAGATTATCAAGGTGGCGATCAAAATTCGGGATCTCTACCGGAACAACAGTTTAAAGCC	389
Oy	282	AAGATGAGCTAGCATTTGATGAGAAATTTAAGAAGAAATTCATCAGTTCGTATGACCG	341
Db	390	AAGAGAGACTGGTTATTGTGGAGAAAGTTCGGAGAAGAGCTGAACAGACCCTCATGACCA	449
Oy	342	TGCTCAGTTTCCATCAGGTGATTTAACCCTTTGACCGGAATGTGTTATCCAGCTGTAA	401
Db	450	TTGTGAGCTTCATAGAGGTGGAATACACTTCGATAGGAACGTGCTCTCAATATTCCTGTC	509
Oy	402	ATGATGACAGAGATGCTGCACCAAAATATTCACGCGCACCTTCACTGCAAGTCAATG	461
Db	510	ATGATGACAGAGCACTGGGTGCATGAACTGTGTCAACGCGCACTTGACGCCACAGATCCACG	569

Qy	462	GAAGGGTAAATATGCTTTGATCATTTTCAGATTTGGAATTTTGGCTGCTTGATA	521
Db	570	GGGCATCAACACGCTTTTAACTCATTTGGCGATGTGGAGTTCTCTCCACCTCATATA	629
Qy	522	ATCCTTTTGGGATTTTAAACCCCACTTACAAAACATATGATGTATCAACAAATGT	581
Db	630	GTCGTGATGGAAGCTGTAGGCCCACTCAAGAGGATTGTGAAGGATCAATTAAGTTGC	689
Qy	582	TGGATGAGGAACATATGA	601
Db	690	TAGATGAGAAAGTCTTTTAA	709

RESULT 40	
AJ720336	
LOCUS	AJ720336
DEFINITION	1589 bp mRNA
ACCESSION	linear
VERSION	VRT 30-SEP-2004
KEYWORDS	Gallus gallus mRNA for hypothetical protein, clone 1514.
SOURCE	AJ720336
	AJ720336
	GI:53133331
	ORF1.
	Gallus gallus (chicken)

REFERENCE

AUTHORS

TITLE
Full-length cDNAs from bursal lymphocytes to facilitate gene
function analysis

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

unpublished
2 (bases 1 to 1589)
Caldwell, R. B.
Direct Submission
Submitted (20-MAY-2004) Caldwell R. B., GSF - Forschungszentrum,

FEATURES	Location/Qualifiers
source	1. .1589

BOU

```

/organism="Gallus gallus"
/mol_type="rRNA"
/strain="CB"
/db_xref="taxon:9031"
/clone="15114"
/cell_type="lymphocyte"
/tissue_type="duresa"
/clone_1ib="rikenl"
/dev_stage="2 weeks old"
1..1589
gene

```

```

5'UTR      /locus_tag="RCJMB04_15114"
            <1.140
            /locus_tag="RCJMB04_15114"
CDS         141..701
            /locus_tag="RCJMB04_15114"
            /note="TOR1"
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="CAJ3199.1"

```

ORIGIN

Query Match	42.0%;	Score 252.4;	DB 5;	Length 1589;
Best Local Similarity	65.8%;	Pred. No. 2.6e-52;		
Matches 367;	Conservative 0;	Mismatches 191;	Indels 0;	Gaps 0

Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rivas, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, M., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrelle, R., Soosa, D., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Vales, R., Vera, V., Villanueva, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Unpublished
2 (bases 1 to 256285)

Worley, K.C.

Direct Submission
Submitted (20-MAR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 256285)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Nov 23, 2002 this sequence version replaced gi:23269911.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

COMMENT

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GSWY
Center clone name: CH230-77C4
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 199255 bases at least Q40
Consensus quality: 203424 bases at least Q30
Consensus quality: 206379 bases at least Q20
Estimated insert size: 195822; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 24557: contig of 24557 bp in length
* 24558 24657: gap of unknown length
* 24658 201810: contig of 177153 bp in length
* 201811 201910: gap of unknown length
* 201911 214313: contig of 12403 bp in length

214314 214413: gap of unknown length
* 214414 247000: contig of 32587 bp in length
* 247001 247100: gap of unknown length
* 247101 248513: contig of 1413 bp in length
* 248514 248613: gap of unknown length
* 248614 249886: contig of 1373 bp in length
* 249887 250086: gap of unknown length
* 250087 251490: contig of 1403 bp in length
* 251490 251586: gap of unknown length
* 251590 252936: contig of 1347 bp in length
* 252937 253036: gap of unknown length
* 253037 256285: contig of 3249 bp in length.

FEATURES
source
1..256285
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-77C4"

misc_feature
1..1142
/note="wgs contig"

misc_feature
4086..6215
/note="wgs contig"

misc_feature
24636..25854
/note="wgs contig"

misc_feature
228541..229886
/note="wgs contig"

misc_feature
245270..247000
/note="wgs contig"

ORIGIN
Query Match 41.6%; Score 250.2; DB 2; Length 256285;
Best Local Similarity 64.7%; Pred. No. 1.1e-51;
Matches 372; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

27 CTCCTCCAGTGGCCACAGATGCTTTAATTCAAAACCTGGCGTTGACGACAAAAGA
120612 CTTTACCCATAGATGCTCTCATGTGTAGTCCAAAGTCTCCCTTCAGCCAGAAAG
87 AGATCTGGGTAAATNGGTGTCAAATCATGCGCCACACCTTAATPAGCGACAAAGTA
120552 AGATCTGGCAAGATGACCAAGCAACCTGTCGCAACATGCTGATGACGACACGAGA
147 GTGAGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
120492 GCGAGATCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
207 CAGAGAAATCATCAAGAACTCTCATGAGACATCATCAAGTGGCCATTTCTTTATGAGA
120432 CCCACAAATCATGAAAGATGTAATTAAGTGGCCATCAAAATTTGATTTCTTACCGGA
267 ATATCAGTTTAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
120372 ACAAACAGTTCACTCAAG
327 ACCTTGCTATGACCGGTGATCACTTTCCATCAAGTGGATTAATCACTTTGACCGAAATGTGT
120312 AAGACGCCATGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
387 TATTCAGGCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
120252 TCTCTTAACCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
447 CTGCAAGTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
120192 GCGCCAGAACCAAG
507 TGGCTGCTTGTATATCTTTTGGAAATTTAAACCCCACTTAACAAAATCATATGATGATG
120132 TTTTCACTCTTCAAGTCCGACAGAAATGACAGGCCCAATCTCAAGAGATTTGGAGG
567 GTATCAACAAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
120072 GAATCAATTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
120038

RESULT 43	AC129440	273225 bp	DNA	linear	HTG 13-MAY-2003
LOCUS	AC129440	273225 bp	DNA	linear	HTG 13-MAY-2003
DEFINITION	ACT129440	273225 bp	DNA	linear	HTG 13-MAY-2003
ACCESSION	AC129440	273225 bp	DNA	linear	HTG 13-MAY-2003
VERSION	AC129440.3	273225 bp	DNA	linear	HTG 13-MAY-2003
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.	273225 bp	DNA	linear	HTG 13-MAY-2003
SOURCE	Rattus norvegicus (Norway rat)	273225 bp	DNA	linear	HTG 13-MAY-2003
ORGANISM	Rattus norvegicus	273225 bp	DNA	linear	HTG 13-MAY-2003
REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	273225 bp	DNA	linear	HTG 13-MAY-2003
AUTHORS	Mueny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooke, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burck, P., Burrell, K., Calderon, E., Cadenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carrillo, L., De Anda, C., Dedetich, D., Delgado, O., Denison, S., Desimo, C., Ding, Y., Dimh, H., Diya, K., Diexper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duvall, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georgegeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, Y., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheewa, L., Louiseged, H., Lozador, R. J., Lu, X., Ma, J., Maheshwari, M., Mahdardine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M. P., McKell, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munguia, M., Murphy, M., Neir, L., Nakevish, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokoilehem, O., Okwunonu, G., Olariu, S., Pal, S., Parks, K., Paternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindester, A., Popovic, D., Primus, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regler, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvatsbeyn, A., Sisson, I., Sitter, C. D., Snajd, D., Sneed, A., Sodergren, E., Song, X. Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svalter, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J., Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlecyk, R., Wooden, H., Wylie, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G., and Gibbs, R. A.	273225 bp	DNA	linear	HTG 13-MAY-2003
TITLE	Unpublished	273225 bp	DNA	linear	HTG 13-MAY-2003
JOURNAL	2 (bases 1 to 273225)	273225 bp	DNA	linear	HTG 13-MAY-2003
REFERENCE	Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	273225 bp	DNA	linear	HTG 13-MAY-2003
TITLE	Submitted (30-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA	273225 bp	DNA	linear	HTG 13-MAY-2003

```

REFERENCE
TITLE      Rat Genome Sequencing Consortium.
JOURNAL
COMMENT
3 (bases 1 to 273225)
Direct Submission
Submitted (13-May-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 13, 2003 this sequence version replaced g1:23264293.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GKTY
Center clone name: CH230-112K14
----- Summary Statistics
Assembly program: Atlas 3.0:
Consensus quality: 224959 bases at least Q40
Consensus quality: 227786 bases at least Q30
Consensus quality: 229294 bases at least Q20
Estimated insert size: 236137; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1. 4842: contig of 4842 bp in length
*
* 4843 4942: gap of unknown length
*
* 4943 265126: contig of 260184 bp in length
*
* 265127 265226: gap of unknown length
*
* 265227 266729: contig of 1503 bp in length
*
* 266730 266829: gap of unknown length
*
* 266830 268869: contig of 2040 bp in length
*
* 268870 268969: gap of unknown length
*
* 268970 270127: contig of 1158 bp in length
*
* 270128 270227: gap of unknown length
*
* 270228 271518: contig of 1291 bp in length
*
* 271519 271618: gap of unknown length
*
* 271619 273225: contig of 1607 bp in length.
*
Location/Qualifiers
1. 273225
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-112K14"
/complement(4198..4940)
/note="clone_boundary"
clone_end:17
site:SCORI
end_sequence:BH258362"
4943. 6894
/note="wgs_contig"

```


Query Match	41.6%	Score 250.2	DB 2	Length 273225
Best Local Similarity	64.7%	Pred. No. 1.1e-51		
Matches 372	Conservative 0	Mismatches 203	Indels 0	Gaps 0
OY	27	CTCCCTCCAGTGGCCACAGATGCTCTTAAATCCAAAAA	CCTGGCCCTTCAGGCACAAAAAGA	86
Db	8649	CTTTACCATATAGTCTCTCATGTGTTTACTCCAAAGATCTGGCCCTTCAGGCCAGAA	8708	
OY	87	AGATCTTGGGTAAATGGTGTCCAAATCCATGCGCACCACTTAAATAGACACACAGTA	146	
Db	8709	AGATCTGAGCAAGATATACCCAGCAAAACGTGGCCCAATGCTGATGATACACACAGTA	8768	
OY	147	GTGAGTCTGTGATAGTCTACAGAGTGACAGGAGATACACCCAAAAA	CAAGAGAGG	206
Db	8769	GCGAATCTTTGATATAGCTGTACAAAGTACCCGAAGAACACCCACAACAAAGAGGAG	8828	
OY	207	CAGAGAAATCATCAAGAACTCATCAAGACAGTATCATAGGTGGCCATTTTATAGTA	266	
Db	8829	CCCAACAAATCATGAAAGATGTAAATAAAGTGGCATCAAATTTGTATTTCTTACCGGA	8888	
OY	267	ATATACATTTAATCAAGATGAGTACGATTGATGAGAGAAATTTAAGAAAGTTTCATC	326	
Db	8889	ACAAACATTCAGTACAGAGAGGTTGTGATGTGAGAAATCCGGAAGAACTGAAC	8948	
OY	327	AGCTTGCTATACACCTGTGCTAGTTTCCATCAGTGGATTTATACCTTTGACCGGAATGT	386	
Db	8949	AGACCGCATGACAAATGGTGTGAGCTTCTACGAAGTGGAGATATACCTTTATATGAAATGGC	9008	
OY	387	TATCCAGGCTGTAAATATGACAGAGATGCTGCACAAATCATTCAGCGCCACTCA	446	
Db	9009	TCTCAATCTTCTGATGATGTCAGAGACCTGTGCTACAGACTGTTACAGGACACTTGA	9068	
OY	447	CTGCCAATCATATGAGACGGGTTAAATATGTTGATCATTTTTCAGATTGTGAATTT	506	
Db	9069	CGCCAGAAACCAACGAGCATAAACATATGTTTCAACCACTTGCCTATGTGAATTC	9128	
OY	507	TGGCTGCTTGTATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTATG	566	
Db	9129	TTTCCACTCTCTACAGTCCGACGGAACCTGACGGCCCAATCTCAAGAGATTGCCAAG	9188	
OY	567	GTATCAACAAATGTTGGATGAGAGACATATGA	601	
Db	9189	GAATCAATTAATTTTAATATGACAAATCTCTGA	9223	
RESULT 44				
LOCUS	BC076797	1938 bp	mRNA	linear
DEFINITION	Xenopus laevis MCC83729 protein, mRNA (CDNA clone MGC:83729			
ACCESSION	BC076797			
VERSION	BC076797.1	GI:49899742		
KEYWORDS	MGC.			
ORGANISM	Xenopus laevis (African clawed frog)			
REFERENCE	Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Alschuler,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stappleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schaeetz,T.E., Brownstein,M.J., Ueding,T.B., Yoshizuki,S.,			
AUTHORS				
TITLE	Genetic and genomic tools for Xenopus research: The NIH Xenopus			
JOURNAL	Dev. Dyn. 225 (4), 384-391 (2002)			
PUBMED	12454917			
REFERENCE	2 (bases 1 to 1938)			
AUTHORS				

38.5%	Score 231.2	DB 5	Length 1938
-------	-------------	------	-------------

Best Local Similarity 64.3%; Pred. No. 5,3e-47;
Matches 347; Conservative 0; Mismatches 193; Indels 0; Gaps 0;

```

QY 50 TTTAATTCAAAAAAGCTGGCGTTTCAGGACAAAGAAAGATCTTGAGTAATGATGTC 109
    |||
DB 151 TTTAGCTCAAGAGCTTGTCTGTTCAAGCCAGAGAAAGATTCAGTAAGATGGCA 210
    |||
QY 110 AAATTCATGCCCAACCTTAAATAGACACAAAGTGTAGAGTGTGATGAGCTTAC 189
    |||
DB 211 AAGACCAATGGCCCAATGCTTATGATGACACAGCAAGTAAATATTGATGAGCTTT 270
    |||
QY 170 AGATGACGAGGAGATGACACCAAAACAGAGAGACAGAAATCATCAAGAACTC 229
    |||
DB 271 AAAGTAACCAAGAAATGTAAAGAACAAAGAGGCCCAAGGCTCTGAAGAGACTT 330
    |||
QY 230 ATCAAGACAGTCAATCAAGCTGCGCATCTTTATAGAAATATCAATTAATCAAGATG 289
    |||
DB 331 GTAAAGGTTGACAGTAAAGTGGCATTTCTAGCAAAATTAACAGTTCACTCGAAGAA 380
    |||
QY 290 CTAGCATGATGAGAAATTTAAGAGAAAGTTCAATCAGCTTGCTATGACCTGTCAGT 349
    |||
DB 391 CTGAATATGTCGAAACCTCAGAGAAAGAACTGAATCAGACTGACATGACGAGTCAGT 450
    |||
QY 350 TTCCATCAGGTGATTAATACCTTTGACCGGATGTTATCCAGGCTGTAATGATGC 409
    |||
DB 451 TTCTTGAAGTTAGTACAGTTGATGATAAAGTGTCTTCCGACTGTGACAGATGT 510
    |||
QY 410 AGAGAGATGCTGACCAAAATCATATTCAGCGCCACCTCCTCCAAAGTCACATGAGAGGT 469
    |||
DB 511 AAAACCTCTTCTTATGAAATGTGAGCGGCACTTAACCAAAATCCCACTGCGATC 570
    |||
QY 470 AATTAATGCTTGTATCATCTTTTTCAGATTTGAAATTTTGGCTGCTGTAAATCCTTT 529
    |||
DB 571 GACAGGTTTATTAATCATCTTTTGTGATGTAATTCCTTCTGCTTATACCTTTGA 630
    |||
QY 530 GGAATTTTAAACCCCACTTACAAAAATCTATGATGATGATCAAAAGTTGATGAA 589
    |||
DB 631 GGAATTTACAGCGCTGCTCAAAAGATTTGTGAAGGGGTCAATTAATTAATTTGATGA 690
    |||

```

```

RESULT 45
BC053167      1813 bp      mRNA      linear      VRT 30-JUN-2004
DEFINITION   BC053167      complete cds.
ACCESSION    BC053167      GI:31418810
VERSION      MGC:
KEYWORDS     Danio rerio (zebrafish)
SOURCE       Danio rerio
ORGANISM     Danio rerio
REFERENCE    1 (bases 1 to 1813)
AUTHORS      Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
              Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
              Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heish, F.,
              Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
              Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
              McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
              Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Huiyk, S.W.,
              Villalón, D.K., Muzny, D.M., Sodergren, E., Lu, X., Gibbs, R.A.,
              Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
              Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
              Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
              Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schultz, J., Myers, R.M.,
              Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
              Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
              Generation and initial analysis of more than 15,000 full-length

```

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1813)
Strausberg, R.
Direct Submission
Submitted (02-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Offices, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
Contact: MGC help desk
Email: cgaabs-remail.nih.gov

Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (JML)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schultz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

FEATURES
source
1..1813
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:63960 IMAGE:6791907"
/rissue_type="Kidney, zebrafish"
/clone_id="NCI_CGAP_ZK1d1"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"

gene
1..1813
/gene="zgc:63960"
/note="synonym: MGC63960"
/db_xref="LocustID:393322"
178..786
/gene="zgc:63960"
/codon_start=1
/product="hypothetical protein MGC63960"
/protein_id="AAH53167.1"
/db_xref="GI:31418811"
/db_xref="LocustID:393322"
/translation="MDSGQSBQSEBELSPHESFNSKSLALQAKKILSKMATAVAN
LITDPSSEILDELKYSREYTSKSKAKRIKIDVYKILKIGILYRNHOFPEEWEET
VRFKKNNQAAFMATVSPFEVEYTPRGILSELLECRDLHELVHNLTMSSHRID
HYFNHADVDTLTXGPESEDRNLNRKICDGIKLLDGLT"

CDS
178..786
/gene="zgc:63960"
/codon_start=1
/product="hypothetical protein MGC63960"
/protein_id="AAH53167.1"
/db_xref="GI:31418811"

ORIGIN
Query Match 37.5%; Score 225.4; DB 5; Length 1813;
Best Local Similarity 63.3%; Pred. No. 1.5e-45;
Matches 346; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

QY 44 GATGCTTTAATTCAAAAAAGCTGGCGTTTCAGGACAAAGAAAGATCTTGAGTAATG 103
    |||
DB 229 GAGAGCTTCAATTCAAAAAGTTGGCCCTTCAGGCTCAAAAGAAAGATTTGAGTAATG 288
    |||
QY 104 GTGTCCAATTCATGCGCCCAACCTTAATAGACACACAAGTAGTAGGCTGATGAG 163
    |||
DB 289 GCCACATGCGCGGTGCGAATCTCTTAACAGACACACAGCGAGATTCGAGCGAA 348
    |||
QY 164 CTCTACAGAGTACCGAGAGTACACCCAAACAGAGAGAGAGAGATCATCAAG 223
    |||
DB 349 CTCTACAGAGAGTACAGAAATACACAGAGAGAGAGAGAGAGAGAGATCATCAAA 408
    |||
QY 224 AACCTCATCAAGACAGTATCAAGCTGCGCATTTCTTATAGAAATATCATGTTAATCAA 283
    |||

```


94	36.2	6.0	28626	4	AA28528	Genomic f
95	36.2	6.0	33206	6	ABN95882	Abn95882 Gene #238
96	36	6.0	647	3	AAC54056	AAC54056 Arabidops
97	35.8	6.0	272	4	AA121218	AA121218 Probe #11
98	35.8	6.0	272	4	ABA66298	ABA66298 Human fce
99	35.8	6.0	272	4	AA164648	AA164648 Probe #15
100	35.8	6.0	272	4	ABA48402	ABA48402 Human bre

ALIGNMENTS

RESULT 1
AB079541
ID AB079541 standard; cDNA; 1915 BP.

AC AB079541;

DT 25-NOV-2002 (first entry)

XX Human SCC-S2 protein encoding cDNA.

KW SSC-S2; apoptosis; tumour; cancer; cytostatic; antisense therapy; human;

XX gene; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 134..700

FT /tag= a

/product= "SCC-S2"

PN MO200259337-A1.

PD 01-AUG-2002.

PF 28-JAN-2002; 2002WC-US002212.

PR 26-JAN-2001; 2001US-0264062P.

PA (GEOU) UNIV GEORGETOWN SCHOOL MEDICINE.

PI Kasid UN, Kumar D, Gokhale P, Ahmad I;

DR MPI: 2002-657478/70.

DR P-PSDB; ABB81967.

XX New SCC-S2 polypeptides and nucleic acids encoding them, useful as a
PT target for identifying compounds that modulate cancer progression by
PT inhibiting apoptosis, as a target for detecting cancers, or for
PT immunizing animals.

PS Claim 2; Fig 1; 70pp; English.

XX The invention relates to an anti-apoptotic gene SSC-S2 and encoded
CC protein. The gene is a positive mediator of tumour growth and metastasis
CC in certain cancer types. The SSC-S2 protein can be expressed by standard
CC recombinant methodology. The SSC-S2 polypeptide is useful as a target for
CC identifying compounds that modulate cancer progression by inhibiting
CC apoptosis, as a target for detecting cancers where the polypeptide is
CC overexpressed, e.g. renal and ovarian cancers, and leukemia. The antibody
CC and antisense oligonucleotide can be used to treat cancer and to inhibit
CC cancer cell proliferation and/or metastasis. The present sequence
CC represents a human SSC-S2 protein encoding cDNA

XX Sequence 1915 BP; 610 A; 331 C; 392 G; 582 T; 0 U; 0 Other;

Query Match 100.0%; Score 601; DB 6; Length 1915;

Best Local Similarity 100.0%; Pred. No. 2.2e-167; Indels 0; Gaps 0;

Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 CTTGACGCTCCGGCGCGCTCCGCGACTCTCCGATGGCAGATGCTTTATTCGA 60
|||||

Db	100	CTTGACGCTCCGGCGCGCTCCGCGACTCTCCGATGGCAGATGCTTTATTCGA	159
Qy	61	AAACCTGGCCGCTTTCAGGCAAAAAGAAGTCTTGGGTAAATGGTGTCCAAATCCATGGC	120
Db	160	AAACCTGGCCGCTTTCAGGCAAAAAGAAGTCTTGGGTAAATGGTGTCCAAATCCATGGC	219
Qy	121	CACCACTTTAATAGACGACAAAGTATGAGGTGTCTGATGAGTCTCTACAGATGACCG	180
Db	220	CACCACTTTAATAGACGACAAAGTATGAGGTGTCTGATGAGTCTCTACAGATGACCG	279
Qy	181	GGAGTACACCCAAAAGAGAGAGAGAGAGATCATCAAGAACTTCATCAAGACAGT	240
Db	280	GGAGTACACCCAAAAGAGAGAGAGAGAGATCATCAAGAACTTCATCAAGACAGT	339
Qy	241	CATCAAGCTGGCCATTTCTTTATAGAAATATATAGTTTATCAAGATGAGCTTGAT	300
Db	340	CATCAAGCTGGCCATTTCTTTATAGAAATATATAGTTTATCAAGATGAGCTTGAT	399
Qy	301	GGAGAAATTTAAGAGAAAGTTTCATCAGCTTGTATGACCGGTGTCAAGTTCCATCAGT	360
Db	400	GGAGAAATTTAAGAGAAAGTTTCATCAGCTTGTATGACCGGTGTCAAGTTCCATCAGT	459
Qy	361	GGATTATACCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGAATGACAGAGATGCT	420
Db	460	GGATTATACCTTTGACCGGAATGTGTATCCAGGCTGTTAAATGAATGACAGAGATGCT	519
Qy	421	GGACCAATCATTCAGGCGCCACCTTCACGCGCAGCATGACAGCGGTTAAATGCTCT	480
Db	520	GGACCAATCATTCAGGCGCCACCTTCACGCGCAGCATGACAGCGGTTAAATGCTCT	579
Qy	481	TGATCATTTTTCAGATGTAATTTTGGCTGCTGTATATCTTTTGGAAATTTTAA	540
Db	580	TGATCATTTTTCAGATGTAATTTTGGCTGCTGTATATCTTTTGGAAATTTTAA	639
Qy	541	ACCCCACTTACAAAACTATGTATGTATCAACAAAATGTTGATGAAGAACATATG	600
Db	640	ACCCCACTTACAAAACTATGTATGTATCAACAAAATGTTGATGAAGAACATATG	699
Qy	601	A 601	
Db	700	A 700	

RESULT 2

ACC81105
ID ACC81105 standard; mRNA; 1892 BP.

AC ACC81105;

DT 25-JUL-2003 (first entry)

XX Human TNF-induced protein Gg2-1 mRNA.

XX Human; T lymphocyte activation; T-cell; A-raf-1; TCRP/PTEN2; asthma;

XX immunosuppressive; antiasthmatic; antiallergic; antiinflammatory;

XX lymphocyte activation; lymphocyte migration; cytokine production;

XX cell surface marker expression; antibody production; apoptosis; allergy;

XX antibody proliferation; antibody differentiation; hypersensitivity;

XX Gg2-1.

OS Homo sapiens.

XX WO2003029277-A2.

PD 10-APR-2003.

PF 02-OCT-2002; 2002WC-US031618.

PR 03-OCT-2001; 2001US-0327212P.

PA (RIG-) RIGEL PHARM INC.

CC treatment of cancer, tumours, developmental abnormalities and foetal
 CC deficiencies, blood disorders, diseases of the immune system, autoimmune
 CC diseases, inflammation, allergies, Alzheimer's and cognitive disorders,
 CC schizophrenia, arthritis, asthma, psoriasis, sepsis, skin disorders,
 CC atherosclerosis, diabetes, cardiovascular disorders, kidney disorders,
 CC digestive/endocrine disorders, infections and AIDS. The polypeptides are
 CC also useful for identifying their binding partners. The sequences given
 CC in AAX97907 to AAX97915 and AAY36223 are used in the exemplification of
 CC the present invention

XX Sequence 1943 BP; 629 A; 325 C; 399 G; 586 T; 0 U; 4 Other;

Query Match 97.9%; Score 588.2; DB 2; Length 1943;

Best Local Similarity 99.3%; Pred. No. 1.4e-163; Mismatches 4; Indels 0; Gaps 0;

Matches 590; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GTCCGGCGCGCGTCCGCACTCTCCGATGGCCACAGATGCTTTTAATCCAAAACCTG 67
 DB 93 GTCCGGCGCGCGTCCGCACTCTCCGATGGCCACAGATGCTTTTAATCCAAAACCTG 152

QY 68 GCGGTTACGACCAAAAGAGATCTGGGTAATAATGTCACAAATCCATCCGACACACC 127
 DB 153 GCGGTTACGACCAAAAGAGATCTGGGTAATAATGTCACAAATCCATCCGACACACC 212

QY 128 TTAATAGACGACACAAGTAGTAGTGCTGATGAGCTCTACAGATGACAGGAGATAC 187
 DB 213 TTAATAGACGACACAAGTAGTAGTGCTGATGAGCTCTACAGATGACAGGAGATAC 272

QY 188 ACCCAAAACAGAGAGAGGACAGAGAAATCATCAGAACCTCTCAAGACAGTCATCANG 247
 DB 273 ACCCAAAACAGAGAGAGGACAGAGAAATCATCAGAACCTCTCAAGACAGTCATCANG 332

QY 248 CTGGCCATCTTTATGAGATATATCATGTTAATCAAGATGAGTACATTTATGAGAAA 307
 DB 333 CTGGCCATCTTTATGAGATATATCATGTTAATCAAGATGAGTACATTTATGAGAAA 392

QY 308 TTTAAGAGAAAGTTTCATCAGCTTGCTATGACCGGTGATCATGTTCCATCAGGTGATAT 367
 DB 393 TTTAAGAGAAAGTTTCATCAGCTTGCTATGACCGGTGATCATGTTCCATCAGGTGATAT 452

QY 368 ACCTTTGACCGGAAATGTTTATTCAGGCTGTTAATGAAATGCAGAGATGCTGCACCA 427
 DB 453 ACCTTTGACCGGAAATGTTTATTCAGGCTGTTAATGAAATGCAGAGATGCTGCACCA 512

QY 428 ATCATTCAGCGCACCTCACTGCGCAAGTCACATGACGAGGTTAATATGCTTTGATCAT 487
 DB 513 ATCATTCAGCGCACCTCACTGCGCAAGTCACATGACGAGGTTAATATGCTTTGATCAT 572

QY 488 TTTTCAGATTTGAAATTTTGGCTGCTGTATATCTTTTGGAAATTTTAAACCCAC 547
 DB 573 TTTTCAGATTTGAAATTTTGGCTGCTGTATATCTTTTGGAAATTTTAAACCCAC 632

QY 548 TTAACAAAACCTATGTATGTATGATCAACAAAATGTTGATGAGAGAAACATATGA 601
 DB 633 TTAACAAAACCTATGTATGTATGATCAACAAAATGTTGATGAGAGAAACATATGA 686

RESULT 4
 ADAL1570
 ID ADAL1570 standard; DNA; 1943 BP.
 AC ADAL1570;
 XX
 XX 06-NOV-2003 (first entry)
 DE Human cDNA encoding a novel secreted protein, SEQ ID NO 98.
 XX
 XX cancer; inflammation; immune disorder; neurological disorder;
 KW blood clotting disorder; food additive; food preservative;
 KW storage capability; fat content; nutritional component; ds; gene; human.
 XX Homo sapiens.
 OS

PN US2003055236-A1.

XX 20-MAR-2003.

XX 14-MAR-2002; 2002US-00097065.

XX 18-DEC-1997; 97US-0068006P.

XX 18-DEC-1997; 97US-0068007P.

XX 18-DEC-1997; 97US-0068008P.

XX 18-DEC-1997; 97US-0068003P.

XX 18-DEC-1997; 97US-0068054P.

XX 18-DEC-1997; 97US-0068057P.

XX 18-DEC-1997; 97US-0068064P.

XX 18-DEC-1997; 97US-0070923P.

XX 19-DEC-1997; 97US-0068169P.

XX 19-DEC-1997; 97US-0068367P.

XX 19-DEC-1997; 97US-0068368P.

XX 19-DEC-1997; 97US-0068369P.

XX 17-JUN-1998; 96WO-US027059.

XX 17-JUN-1998; 99US-00334595.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Moore PA, Ruben SM, Carter KC, Shi Y, Rosen CA, Soppet DR;

XX Kyaw H, Wei Y, Florence KA, Duan DR, Florence C, Greene JM, Feng P;

XX Ferrie AM, Yu G, Janat F, Ni J;

XX WPI; 2003-567105/53.

XX P-PSDB; ADAL1694.

XX New secreted HKABT24 nucleic acid molecules and polypeptides, useful for

XX preventing, treating, or ameliorating a medical condition, such as

XX cancer, inflammation, immune disorders, neurological and blood clotting

XX disorders.

XX Claim 1; SEQ ID NO 98; 118pp; English.

XX The invention relates to an isolated HKABT24 nucleic acid molecule. The

XX polypeptides, nucleic acids and antibodies are useful for diagnosing a

XX pathological condition or a susceptibility to a pathological condition,

XX for preventing, treating, or ameliorating a medical condition, such as

XX cancer, inflammation and other immune disorders, neurological and blood

XX clotting disorders. The nucleic acids are also useful for chromosome

XX identification, radiation hybrid mapping or long-range restriction

XX mapping. The polypeptides and antibodies are useful for providing

XX immunological probes for differential identification of the tissues

XX immunohistochemistry assays. The polypeptide, polynucleotide, agonist or

XX antagonist may also be used as a food additive or preservative to

XX increase or decrease storage capabilities, fat content or other

XX nutritional components. The present sequence represents cDNA encoding a

XX novel human secreted protein. Note: The sequence data for this patent did

XX not form part of the printed specification but was obtained in electronic

XX CC format directly from USPTO at

XX segdata.uspto.gov.uk/sequence.html?DocID=20030055236.

XX Sequence 1943 BP; 629 A; 325 C; 399 G; 586 T; 0 U; 4 Other;

Query Match 97.9%; Score 588.2; DB 9; Length 1943;

Best Local Similarity 99.3%; Pred. No. 1.4e-163; Mismatches 4; Indels 0; Gaps 0;

Matches 590; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 8 GTCCGGCGCGCGTCCGCACTCTCCGATGGCCACAGATGCTTTAATCCAAAACCTG 67
 DB 93 GTCCGGCGCGCGTCCGCACTCTCCGATGGCCACAGATGCTTTAATCCAAAACCTG 152

QY 68 GCGGTTACGACCAAAAGAGATCTGGGTAATAATGTCACAAATCCATCCGACACACC 127
 DB 153 GCGGTTACGACCAAAAGAGATCTGGGTAATAATGTCACAAATCCATCCGACACACC 212

QY 128 TTAATAGACGACACAAGTAGTAGTGCTGATGAGCTCTACAGATGACAGGAGATAC 187
 DB 213 TTAATAGACGACACAAGTAGTAGTGCTGATGAGCTCTACAGATGACAGGAGATAC 272

PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX
 PI Oca T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Makamatsu A, Nagai K, Otsuki T,
 XX MPI, 2001-318749/34.
 DR
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 8; SEQ ID NO 12578; 2537bp + Sequence listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 CC
 SQ Sequence 1729 BP; 563 A; 277 C; 354 G; 535 T; 0 U; 0 Other;

Query Match 93.9%; Score 564.4; DB 4; Length 1729;
 Best Local Similarity 99.8%; Pred. No. 1.5e-156;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAAATTCGAAACCTGGCCGTTGAGCAGCAAAAGAGATCTTG 95
 DB 102 TGGCCACAGATGCTTTAAATTCGAAACCTGGCCGTTGAGCAGCAAAAGAGATCTTG 161
 QY 96 GTTAAATGCTGTCGAAATTCGACCACTTAATAGACAGACAGATGAGGTGC 155
 DB 162 GTTAAATGCTGTCGAAATTCGACCACTTAATAGACAGACAGATGAGGTGC 221
 QY 156 TGGATAGCTCTTACAGAGTACCCAGGAGTACCCGAAACAGAGAGGAGGAGAGAA 215
 DB 222 TGGATAGCTCTTACAGAGTACCCAGGAGTACCCGAAACAGAGAGGAGGAGAGAA 281
 QY 216 TCATCAAGACCTCATCAAGACAGTCAAGCTGGCATTCTTTATAGGAATATCACT 275
 DB 282 TCATCAAGACCTCATCAAGACAGTCAAGCTGGCATTCTTTATAGGAATATCACT 341
 QY 276 TTATATCAAGATGAGTCAAGATGAGAGAAATTTAAGAGAAAGTTCACTGCTCTA 335
 DB 342 TTATATCAAGATGAGTCAAGATGAGAGAAATTTAAGAGAAAGTTCACTGCTCTA 401
 QY 336 TGAACGTGCTGCTTCCATCAGTGAATATACCTTTACCGGAATGTTATCCAGGC 395
 DB 402 TGAACGTGCTGCTTCCATCAGTGAATATACCTTTACCGGAATGTTATCCAGGC 461
 QY 396 TGTAAATGATGACAGAGATGCTCAACCAATCACTTACGCCCACTCACTGCCAGT 455
 DB 462 TGTAAATGATGACAGAGATGCTCAACCAATCACTTACGCCCACTCACTGCCAGT 521

QY 456 CACATGACGGGTTATATGCTTGATCAATTTCAATTTGTAATTTGGCGCT 515
 DB 522 CACATGACGGGTTATATGCTTGATCAATTTCAATTTGTAATTTGGCGCT 581
 QY 516 TGTATATCTTTTGGAAATTTAAACCCCACTTACAAAACTATGATGATCAACA 575
 DB 582 TGTATATCTTTTGGAAATTTAAACCCCACTTACAAAACTATGATGATCAACA 641
 QY 576 AAATGTTGATGAGAGAAATATGTA 601
 DB 642 AAATGTTGATGAGAGAAATATGTA 667

RESULT 8
 ID ADR14214
 XX ADR14214 standard; DNA; 1814 BP.
 XX
 AC ADR14214;
 XX
 DT 21-OCT-2004 (first entry)
 XX
 DE Human NF-kappaB pathway-associated gene SeqID215.
 XX
 KW NF-kappaB pathway; antiinflammatory; cytosolic; hepatotropic; virulicide;
 KW antitartaric; antirheumatic; gastrointestinal-gen; antiaesthetic;
 KW antilactobacilic; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnery; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; gene; ds; human.
 XX
 OS Homo sapiens.
 XX
 XX
 PN WO2004065577-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 13-JAN-2004; 2004WO-US000798.
 XX
 PR 14-JAN-2003; 2003US-0440068P.
 XX
 PR 12-MAY-2003; 2003US-0469757P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Nadler SG, Neubauer MG, Feder JN, Carman J;
 XX MPI; 2004-562168/54.
 DR P-PSDB; ADR14215.
 XX
 PT New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 PS Claim 1; SEQ ID NO 215; 237bp; English.
 XX
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytostatic, hepatotropic, virulicide, antitartaric, antirheumatic,
 CC gastrointestinal-gen, antiaesthetic, antilactobacilic,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnery activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions

or diseases associated with the NF-kappaB pathway. The condition is an immune disorder, an inflammatory disorder, cancer, aberrant apoptosis, CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis, CC hepatic disorders, Hodgkin's lymphoma, haematopoietic tumours, hyper-IgM CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper CC immune activity, disorders related to aberrant acute phase responses, CC hypercongenital conditions, birth defects, necrotic lesions, wounds, CC organ transplant rejection, conditions related to organ transplant CC rejection, disorders related to aberrant signal transduction, CC proliferating disorders, cancers and HIV propagation in cells infected CC with other viruses. The present sequence is that of a human gene which is CC subject to the novel association with the NF-kappaB pathway of the CC invention. Note: This sequence does not appear in the specification but CC was obtained by the indexer from Genbank.

Sequence 1814 BP; 590 A; 297 C; 372 G; 555 T; 0 U; 0 Other;

Query Match 93.9%; Score 564.4; DB 13; Length 1814;
Best Local Similarity 99.8%; Pred. No. 1.6e-156;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCGGTGACGACAAAGAAATCTTGG 95
175 TGGCCACAGATGCTTTAATTCAGAAACCTGGCGGTGACGACAAAGAAATCTTGG 234
96 GTAAATGGTGTCCAAATCCATCGCACCACTTAATAGACACAAAGTAGAGGTGC 155
235 GTAAATGGTGTCCAAATCCATCGCACCACTTAATAGACACAAAGTAGAGGTGC 294
156 TGGATAGCTCTACAGAGTACCAAGAGTACCAAGAGAGAGAGAGAGAGAGAG 215
295 TGGACGAGCTCTACAGAGTACCAAGAGTACCAAGAGAGAGAGAGAGAGAGAG 354
216 TCATCAAGAACCTCATCAAGAGTACCAAGTGGCGCTTTTATAGAAATCAAGT 275
355 TCATCAAGAACCTCATCAAGAGTACCAAGTGGCGCTTTTATAGAAATCAAGT 414
276 TTAATCAAGTAGAGTACGATGATGAGAAATTTAAGAAAGATTCACCTTGCTA 335
415 TTAATCAAGTAGAGTACGATGATGAGAAATTTAAGAAAGATTCACCTTGCTA 474
336 TGAACGCTGCTCACTTCATCAGGTGATTAATCTTGAACGGAATGTTATCAAGC 395
475 TGAACGCTGCTCACTTCATCAGGTGATTAATCTTGAACGGAATGTTATCAAGC 534
396 TGTAAATGAATGACAGAGATGCTGACCAATCACTTCAAGGCGCCACTCAGCAAGT 455
535 TGTAAATGAATGACAGAGATGCTGACCAATCACTTCAAGGCGCCACTCAGCAAGT 594
456 CACATGAGAGCGGTTAATTAATGCTTTGATCATTTTTCAGATTTTGGTGCCT 515
595 CACATGAGAGCGGTTAATTAATGCTTTGATCATTTTTCAGATTTTGGTGCCT 654
516 TGTATATCTTTTGGGAATTTTAAACCCCACTTAACAAAATATGATGATGATCAAC 575
655 TGTATATCTTTTGGGAATTTTAAACCCCACTTAACAAAATATGATGATGATCAAC 714
576 AATGTTGATGAAGAAACATATGA 601
715 AATGTTGATGAAGAAACATATGA 740
```

RESULT 9
ADRI4216
ID ADRI4216 standard; DNA; 2003 BP.
XX AC ADRI4216;
XX

DT 21-OCT-2004 (first entry)

DE Human NF-kappaB pathway-associated gene SegID217.

XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
XX antiarthritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
XX antiatherosclerotic; immunomodulator; cerebroprotective; vasotrophic;
XX immunosuppressive; vulnary; gene therapy; immune disorder;
XX inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
XX hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
XX hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
XX X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
XX viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
XX viral replication; host cell survival; evasion of immune response;
XX rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
XX atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
XX autoimmune disorder; hyper immune activity;
XX aberrant acute phase response; hypercongenital condition; birth defect;
XX necrotic lesion; wound; organ transplant rejection;
XX aberrant signal transduction; proliferating disorder; cancer;
XX HIV propagation; gene; ds; human.

XX Homo sapiens.

XX MO2004065577-A2.

XX 05-AUG-2004.

XX 13-JUN-2004; 2004MO-US000798.

XX 14-JAN-2003; 2003US-0440068P.

XX 12-MAY-2003; 2003US-0469757P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Nadler SG, Neubauer MG, Feder JN, Carman J;

XX MPI; 2004-562168/54.

XX P-PSDB; ADRI4217.

XX New isolated polynucleotides and polypeptides associated with NF-kappaB
XX pathway, useful for diagnosing, treating, or preventing disorders or
XX diseases associated with NF-kappaB pathway.

XX Claim 1; SEQ ID NO 217; 237bp; English.

XX This invention relates to the novel association of protein sequences (and
XX the genes which encode them) to the NF-kappaB pathway. The invention may
XX be useful for the production of compounds with an antiinflammatory,
XX cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
XX gastrointestinal-gen, antiasthmatic, antiatherosclerotic,
XX immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
XX vulnary activity or for gene therapy. The proteins and nucleotides are
XX useful for diagnosing, preventing, treating, or ameliorating conditions
XX of diseases associated with the NF-kappaB pathway. The condition is an
XX immune disorder, an inflammatory disorder, an inflammatory disorder
XX related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
XX hepatic disorders, Hodgkin's lymphoma, haematopoietic tumours, hyper-IgM
XX syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
XX ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
XX hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
XX survival, evasion of immune responses, rheumatoid arthritis, inflammatory
XX bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
XX syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
XX immune activity, disorders related to aberrant acute phase responses,
XX hypercongenital conditions, birth defects, necrotic lesions, wounds,
XX organ transplant rejection, conditions related to organ transplant
XX rejection, disorders related to aberrant signal transduction,
XX proliferating disorders, cancers and HIV propagation in cells infected
XX with other viruses. The present sequence is that of a human gene which is
XX subject to the novel association with the NF-kappaB pathway of the
XX invention. Note: This sequence does not appear in the specification but
XX was obtained by the indexer from Genbank.

XX Sequence 2003 BP; 667 A; 324 C; 406 G; 606 T; 0 U; 0 Other;

Query Match 93.9%; Score 564.4; DB 13; Length 2003;

Best Local Similarity 99.8%; Pred. No. 1.6e-156;

Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTAAATTCGAAAACCTGGCCCTTCGCGCAAAAAGAAATCTTGG 95
 DB 205 TGGCCACAGATGCTTAAATTCGAAAACCTGGCCCTTCGCGCAAAAAGAAATCTTGG 264
 QY 96 GTAAAAATGCTGTCGCAATCCATCGCCGACACCTTAATAGACACAAAGTAGTGGTGC 155
 DB 265 GTAAAAATGCTGTCGCAATCCATCGCCGACACCTTAATAGACACAAAGTAGTGGTGC 324
 QY 156 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAAGAGGACAGAGA 215
 DB 325 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAAGAGGAGAGAGA 384
 QY 216 TCATCAAGAACCTCATCAAGACAGTCACTCAAGCTGGCATTTCTTATAGAAATATCA 275
 DB 385 TCATCAAGAACCTCATCAAGACAGTCACTCAAGCTGGCATTTCTTATAGAAATATCA 444
 QY 276 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCATCAGCTGCTA 335
 DB 445 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCATCAGCTGCTA 504
 QY 336 TGAACGCTGTCAGTTTCATCAGGTGATTAATACCTTTGACCGAATGTTTATCCAGGC 395
 DB 505 TGAACGCTGTCAGTTTCATCAGGTGATTAATACCTTTGACCGAATGTTTATCCAGGC 564
 QY 396 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCATCGCAAGT 455
 DB 565 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCATCGCAAGT 624
 QY 456 CACATGAGAGGGTTAATATGCTTTGATGATTTTTCAGATGTTGAAATTTTGGCTGCTC 515
 DB 625 CACATGAGAGGGTTAATATGCTTTGATGATTTTTCAGATGTTGAAATTTTGGCTGCTC 684
 QY 516 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGATGATATCAACA 575
 DB 685 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGATGATATCAACA 744
 QY 576 AAATGTTGATGAAGAGAAATATGA 601
 DB 745 AAATGTTGATGAAGAGAAATATGA 770
 RESULT 10
 ACN44179
 ID ACN44179 standard; cDNA; 2034 BP.
 ACN44179;
 XX 18-NOV-2004 (first entry)
 DE Human mRNA sequence HCT1950131.
 KM Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.
 OS Homo sapiens.
 XX WO2003073826-A2.
 PD 12-SEP-2003.
 PF 28-FEB-2003; 2003MO-US006235.
 PR 01-MAR-2002; 2002US-00087192.
 PA (SAGR-) SAGRES DISCOVERY.
 PI Morris DW;

XX WPI; 2003-328604/31.

DR Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

PT comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 497; Opp; English.

CC The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to CarcinoMA Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining CarcinoMA Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA gene
 CC sequence. Note: This patent is an equivalent to basic patent
 CC US2002182586A1, for which no sequence data was published

XX Sequence 2034 BP; 639 A; 352 C; 412 G; 631 T; 0 U; 0 Other;

Query Match 93.9%; Score 564.4; DB 11; Length 2034;

Best Local Similarity 99.8%; Pred. No. 1.7e-156;

Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTAAATTCGAAAACCTGGCCCTTCGCGCAAAAAGAAATCTTGG 95
 DB 249 TGGCCACAGATGCTTAAATTCGAAAACCTGGCCCTTCGCGCAAAAAGAAATCTTGG 308
 QY 96 GTAAAAATGCTGTCGCAATCCATCGCCGACACCTTAATAGACACAAAGTAGTGGTGC 155
 DB 309 GTAAAAATGCTGTCGCAATCCATCGCCGACACCTTAATAGACACAAAGTAGTGGTGC 368
 QY 156 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAAGAGGACAGAGA 215
 DB 369 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAAAGAGGACAGAGA 428
 QY 216 TCATCAAGAACCTCATCAAGACAGTCACTCAAGTGGCCATTTCTTATAGAAATATCA 275
 DB 429 TCATCAAGAACCTCATCAAGACAGTCACTCAAGTGGCCATTTCTTATAGAAATATCA 488
 QY 276 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCATCAGCTGCTA 335
 DB 489 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCATCAGCTGCTA 548
 QY 336 TGAACGCTGTCAGTTTCATCAGGTGATTAATACCTTTGACCGAATGTTTATCCAGGC 395
 DB 549 TGAACGCTGTCAGTTTCATCAGGTGATTAATACCTTTGACCGAATGTTTATCCAGGC 608
 QY 396 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGGCGCCACCTCATCGCAAGT 455
 DB 609 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGGCGCCACCTCATCGCAAGT 668
 QY 456 CACATGAGAGGGTTAATATGCTTTGATGATTTTTCAGATGTTGAAATTTTGGCTGCTC 515
 DB 669 CACATGAGAGGGTTAATATGCTTTGATGATTTTTCAGATGTTGAAATTTTGGCTGCTC 728
 QY 516 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGATGATATCAACA 575
 DB 729 TGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATCATGATGATATCAACA 788
 QY 576 AAATGTTGATGAAGAGAAATATGA 601
 DB 789 AAATGTTGATGAAGAGAAATATGA 814
 RESULT 11
 ADRI4038
 ID ADRI4038 standard; DNA; 2081 BP.

XX ADR14038;
 XX
 XX 21-OCT-2004 (first entry)
 XX
 DE Human NF-kappaB pathway-associated gene Segid39.
 XX
 XX NF-kappaB pathway; antiinflammatory; cytostatic; hepatotropic; virucide;
 KW antirheumatic; antirheumatic; gastrointestinal-gen; antiaesthetic;
 KW antiarteriosclerotic; immunomodulator; cerebroprotective; vasotropic;
 KW immunosuppressive; vulnery; gene therapy; immune disorder;
 KW inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KW hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KW hyper-IgM syndrome; hypohidrotic ectodermal dysplasia;
 KW X-linked anhidrotic ectodermal dysplasia; immunodeficiency;
 KW viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
 KW viral replication; host cell survival; evasion of immune response;
 KW rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KW autoimmune disorder; hyper immune activity;
 KW aberrant acute phase response; hypercongenital condition; birth defect;
 KW necrotic lesion; wound; organ transplant rejection;
 KW aberrant signal transduction; proliferating disorder; cancer;
 KW HIV propagation; gene; ds; human.
 XX
 OS Homo sapiens.
 XX
 PN MO2004065577-A2.
 XX
 PD 05-AUG-2004.
 XX
 PF 13-JAN-2004; 2004MO-US000798.
 XX
 PR 14-JAN-2003; 2003US-0440068P.
 XX
 PR 12-MAY-2003; 2003US-0469757P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX
 PI Nadler SG, Neubauer MG, Feder JN, Carman J;
 XX
 XX WPI; 2004-562168/54.
 DR
 DR P-PSDB; ADR14039.
 PT
 PT New isolated polynucleotides and polypeptides associated with NF-kappaB
 PT pathway, useful for diagnosing, treating, or preventing disorders or
 PT diseases associated with NF-kappaB pathway.
 XX
 XX Claim 1; SEQ ID NO 39; 237pp; English.
 PS
 CC This invention relates to the novel association of protein sequences (and
 CC the genes which encode them) to the NF-kappaB pathway. The invention may
 CC be useful for the production of compounds with an antiinflammatory,
 CC cytostatic, hepatotropic, virucide, antirheumatic, antipneumatic,
 CC gastrointestinal-gen, antiaesthetic, antiarteriosclerotic,
 CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
 CC vulnery activity or for gene therapy. The proteins and nucleotides are
 CC useful for diagnosing, preventing, treating, or ameliorating conditions
 CC or diseases associated with the NF-kappaB pathway. The condition is an
 CC immune disorder, an inflammatory disorder, cancer, aberrant apoptosis,
 CC related to aberrant NF-kappaB regulation, cancer, aberrant apoptosis,
 CC hepatic disorders, Hodgkin's lymphoma, haematopoietic tumours, hyper-IgM
 CC syndromes, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
 CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
 CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
 CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
 CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
 CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
 CC immune activity, disorders related to aberrant acute phase responses,
 CC hypercongenital conditions, birth defects, necrotic lesions, wounds,
 CC organ transplant rejection, conditions related to organ transplant
 CC rejection, disorders related to aberrant signal transduction,
 CC proliferating disorders, cancers and HIV propagation in cells infected
 CC with other viruses. The present sequence is that of a human gene which is

CC subject to the novel association with the NF-kappaB pathway of the
 CC invention. Note: This sequence does not appear in the specification but
 CC was obtained by the indexer from Genbank.

XX SQ Sequence 2081 BP; 668 A; 359 C; 422 G; 632 T; 0 U; 0 Other;

Query Match 93.9%; Score 564.4; DB 13; Length 2081;
 Best Local Similarity 99.8%; Pred. No. 1.7e-156;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTTAATTCGAAAACCTGGCGTTCAGGCACAAAAAGAGATCTTGG 95
 DB 268 TGGCCACAGATGCTTTTAATTCGAAAACCTGGCGTTCAGGCACAAAAAGAGATCTTGG 327
 QY 96 GTAAATAGTGTGCCAAATCCATGCGACACCTTAATGAGACCAAGTAGTGGTGC 155
 DB 328 GTAAATAGTGTGCCAAATCCATGCGACACCTTAATGAGACCAAGTAGTGGTGC 387
 QY 156 TGGATGAGCTCTACAGAGTGAACAGGAGTACACCCAAAACAAGAGAGCAGAGAGA 215
 DB 388 TGGATGAGCTCTACAGAGTGAACAGGAGTACACCCAAAACAAGAGAGCAGAGAGA 447
 QY 216 TCATCAAGAACTCATCAAGACAGTCATCAGCTGCCCATCTTTATAGAAATATCACT 275
 DB 448 TCATCAAGAACTCATCAAGACAGTCATCAGCTGCCCATCTTTATAGAAATATCACT 507
 QY 276 TTAATCAAGATAGCTAGCATGATGAGGAATTTAAGAAAGATTCATCGCTTGCTA 335
 DB 508 TTAATCAAGATAGCTAGCATGATGAGGAATTTAAGAAAGATTCATCGCTTGCTA 567
 QY 336 TGACGTGTGTCAGTTTCCATCAGTGTGATTTATACCTTTGACCGGATGTGTTATCCAGGC 395
 DB 568 TGACGTGTGTCAGTTTCCATCAGTGTGATTTATACCTTTGACCGGATGTGTTATCCAGGC 627
 QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCCACCTCACTGCCAGT 455
 DB 628 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCCACCTCACTGCCAGT 687
 QY 456 CACATGAGACGGGTAAATATGCTTGATCATTTTTCAGATGTGTAATTTTGGCGCT 515
 DB 688 CACATGAGACGGGTAAATATGCTTGATCATTTTTCAGATGTGTAATTTTGGCGCT 747
 QY 516 TGTAAATGCTTTTGGGAATTTAAACCCCACTTCAAAAACATATGTGATGATCAACA 575
 DB 748 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTCAAAAACATATGTGATGATCAACA 807
 QY 576 AATGTTGATGAAGAGAAACATATGA 601
 DB 808 AATGTTGATGAAGAGAAACATATGA 833

RESULT 12
 ACN44178
 ID ACN44178 standard; DNA; 58723 BP.

XX ACN44178;

DT 18-NOV-2004 (first entry)

DE Human genomic sequence hCG36837.

KW Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.

OS Homo sapiens.

PN WO2003073826-A2.

PD 12-SEP-2003.

PF 28-FEB-2003; 2003WO-US006235.

PR 01-MAR-2002; 2002US-00087192.

PA (SAGR-) SAGRES DISCOVERY.
 XX
 PI Morris DW;
 XX
 DR WPI; 2003-328604/31.
 XX
 PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.
 XX
 XX Claim 1; SEQ ID NO 496; Opp; English.
 XX
 XX The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcino Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcino Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA coding
 CC sequence. Note: This patent is an equivalent to basic patent
 CC US2002182586A1, for which no sequence data was published
 XX
 XX Sequence 58723 BP; 16691 A; 11477 C; 12539 G; 17816 T; 0 U; 0 Other;
 SQ
 Query Match 93.9%; Score 564.4; DB 11; Length 58723;
 Best Local Similarity 99.8%; Pred. No. 8.9e-156;
 Matches 56; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 XX
 QY 36 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 95
 DB 46938 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 46997
 QY 96 GTAAATGCTGCTCAATCATCGCCACCTTAATAGACGACACAAAGTAGAGGTGC 155
 DB 46998 GTAAATGCTGCTCAATCATCGCCACCTTAATAGACGACACAAAGTAGAGGTGC 47057
 QY 156 TGGATGAGCTCTACAGAGTACAGGAGGATACACCCAAAACAAGAGAGGACAGAGA 215
 DB 47058 TGGATGAGCTCTACAGAGTACAGGAGGATACACCCAAAACAAGAGAGGACAGAGA 47117
 QY 216 TCATCAAGAACTCATCAAGACAGCTATCAAGCTGGCCATTCTTTATAGGAATATCAGT 275
 DB 47118 TCATCAAGAACTCATCAAGACAGCTATCAAGCTGGCCATTCTTTATAGGAATATCAGT 47177
 QY 276 TTATCAAGATGAGCTTACGCTTATGAGAAATTTAAGAAAGATTATCAGCTTCTTA 335
 DB 47178 TTATCAAGATGAGCTTACGCTTATGAGAAATTTAAGAAAGATTATCAGCTTCTTA 47237
 QY 336 TGACCGTGTCACTTCCATCAGGTGATATACCTTTGACCGGAATGTTTATCCAGAGC 395
 DB 47238 TGACCGTGTCACTTCCATCAGGTGATATACCTTTGACCGGAATGTTTATCCAGAGC 47297
 QY 396 TGTAAATGAAATCAGAGAGATCTGACCAAAATCATTCAGCGCCACCTCATGCCAAGT 455
 DB 47298 TGTAAATGAAATCAGAGAGATCTGACCAAAATCATTCAGCGCCACCTCATGCCAAGT 47357
 QY 456 CACATGAGCGGTTAATATATGCTTGTATCATTTTTCAGATTGTAATTTTGGCTGCTT 515
 DB 47358 CACATGAGCGGTTAATATATGCTTGTATCATTTTTCAGATTGTAATTTTGGCTGCTT 47417
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATATCAACA 575
 DB 47418 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATATCAACA 47477
 QY 576 AAATGTTGATGAAGAGACATATGA 601
 DB 47478 AAATGTTGATGAAGAGACATATGA 47503

RESULT 13
 ID AAH07403
 XX AAH07403 standard; cDNA; 816 BP.
 ID
 AC AAH07403;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA clone (5'-primer) SEQ ID NO:4238.
 XX
 KM Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 XX
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.
 PA
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Iishi S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 XX Claim 1; SEQ ID NO 4238; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialized methods. AAH03166 to AAH13628 and
 CC AAH1633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 816 BP; 262 A; 155 C; 180 G; 212 T; 0 U; 7 Other;
 Query Match 91.1%; Score 547.6; DB 4; Length 816;
 Best Local Similarity 98.9%; Pred. No. 1e-151;
 Matches 56; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
 XX
 QY 36 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 95
 DB 102 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTCCAGGACAAAAGAGATCTTGG 161

OY	96	GTAAATAGTGTCCAAATCCATGCGCCACCACTTAATAGCAGACAAAGTAGAGGTGC	155
Db	162	GTAAATAGTGTCCAAATCCATGCGCCACCACTTAATAGCAGACAAAGTAGAGGTGC	221
OY	156	TGGATGAGCTCTCAAGAGTGACCAAGGAGTACCCCAAAACAAGAGAGGCAGAGAAGA	215
Db	222	TGGATGAGCTCTCAAGAGTGACCAAGGAGTACCCCAAAACAAGAGAGGCAGAGAAGA	281
OY	216	TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTAATAGGAATTAATCAGT	275
Db	282	TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTAATAGGAATTAATCAGT	341
OY	276	TTAATCAAGATGAGCTGACATTGATGAGAGAAATTTAAGAGAAAGTTTCATCAGCTTGCTA	335
Db	342	TTAATCAAGATGAGCTGACATTGATGAGAGAAATTTAAGAGAAAGTTTCATCAGCTTGCTA	401
OY	336	TGACCGTGTGCACTTTCCATCAGTGTGATTAATCCTTTGACCGGAATGTGTATTCAGAGC	395
Db	402	TGACCGTGTGCACTTTCCATCAGTGTGATTAATCCTTTGACCGGAATGTGTATTCAGAGC	461
OY	396	TGTTAAATGAAATGACAGAGAGATGCTGSCAACAAATCATTAAGCGCCCACTCACTGCCAAGT	455
Db	462	TGTTAAATGAAATGACAGAGAGATGCTGSCAACAAATCATTAAGCGCCCACTCACTGCCAAGT	521
OY	456	CACATGGAAGGGTTAATATGTCTTTGATCATTTTTCAAGTTGTGAATTTTTGGCTGCCCT	515
Db	522	CACATGGAAGGGTTAATATGTCTTTGATCATTTTTCAAGTTGTGAATTTTTGGCTGCCCT	581
OY	516	TGTATTAATCCTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATGATACACA	575
Db	582	TGTATTAATCCTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATGATACACA	641
OY	576	AAAT--GTTGATGAAGAGACATATGA	601
Db	642	AAATGTTTGGATGGAAGAACATATTA	669

```

RESULT 14
ACH89719
ID ACH89719 standard; DNA; 544 BP.
XX
XX ACH89719;
AC
XX
XX 29-JUL-2004 (first entry)
XX
XX Human genome derived single exon probe #22914.
DE
XX
XX Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX
XX Homo sapiens.
OS
XX
XX US2003194704-A1.
PN
XX
XX 16-OCT-2003.
PD
XX
XX 03-APR-2002; 2002US-00029386.
PF
XX
XX 03-APR-2002; 2002US-00029386.
PR
XX
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
XX Penn SG, Rank DR, Hanzel DK;
PI
XX
XX WPI; 2004-119264/12.
DR
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX

```

PS Claim 1; SEQ ID NO 22914; 80pp; English.

The invention relates to a nucleic acid probe for measuring human gene expression, comprising any of the 27,400 fully defined nucleotide sequences in the specification, or their complements or fragments, and encoding at least 8 amino acids of any of the 6888 amino acid sequences fully defined in the specification. The probe is a single exon probe that hybridizes under high stringency conditions to a nucleic acid molecule expressed in human cells or tissues. Also included are a spatially-addressable set of single exon nucleic acid probes for measuring human gene expression (comprising a plurality of single exon nucleic acid probes cited above, where each of the plurality of probes is separately and addressably isolatable or amplifiable from the plurality), a single exon microarray for measuring human gene expression, a method of measuring human gene expression, a vector comprising the single exon probe cited above, an ORF-encoded peptide comprising at least 8 contiguous amino acids of any of the above-mentioned amino acid sequences (optionally with conservative amino acid substitutions), an isolated antibody that binds specifically to a peptide cited above, methods of selling and/or licensing single exon probes or microarrays to a customer desirous to measure gene expression, a method of providing human gene expression data by subscrption, and a computer-readable storage medium which contains a database having a plurality of records (each record including data on the expression of a single exon probe cited above. The probe, methods and apparatus are useful in gene expression analysis. The probes may be used as tools for surveying tissues to detect the presence of expressed messages that contain their specific exon, or in constructing genome-derived single exon microarrays. In addition, the probes are used in identifying and characterizing alternative splicing events, in detecting and characterizing gross alterations in the genomic locus that includes their exon, in assessing smaller genomic alterations, in priming the synthesis of nucleic acids, or in expressing the ORF-encoded peptide. The present sequence is a human single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at seqdata.uspto.gov/sequence.html?DocID=20030194704

SQ Sequence 544 BP; 175 A; 109 C; 117 G; 143 T; 0 U; 0 Other;

Query Match 90.2%; Score 542.4; DB 12; Length 544;

Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0.

OY	30	GGGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTTACGGGACCAAAAAGAGATCTGG	95
Db	1	TGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTTACGGGACCAAAAAGAGATCTGG	60
OY	96	GTAATAATGGTGTCCAAATCCATCGCCACACCTTAATAGACACAGTAGTAGGTGC	155
Db	61	GTAATAATGGTGTCCAAATCCATCGCCACACCTTAATAGACACAGTAGTAGGTGC	120
OY	156	TGGATGAGCTCTACAGAGTGACCAAGGAGTACACCCAAAACAGAGAGAGCAGAGAGA	215
Db	121	TGGATGAGCTCTACAGAGTGACCAAGGAGTACACCCAAAACAGAGAGAGCAGAGAGA	180
OY	216	TCATCAAGAAACCTCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATPATCAGT	275
Db	181	TCATCAAGAAACCTCAATCAAGACAGTCATCAAGCTGGCCATCTTTATAGAAATPATCAGT	240
OY	276	TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGAAATTTCATCAGCTTGTA	335
Db	241	TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGAAATTTCATCAGCTTGTA	300
OY	336	TGACCTGGTCAAGTTTCCATCAGGTGATTTACCTTTGACCGGAATGTGTATCCAGGC	395
Db	301	TGACCTGGTCAAGTTTCCATCAGGTGATTTACCTTTGACCGGAATGTGTATCCAGGC	360
OY	396	TGTTAATGAATGACAGAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTCCCAAGT	455
Db	361	TGTTAATGAATGACAGAGAGATGCTGCACCAATCATTCAGCGCCACCTCAGTCCCAAGT	420
OY	456	CACATGAGCGGTTAATAATGCTTTGATCAATTTTTAGATGTGTGAATTTTTGGCTGCCCT	515

Db 421 CACATGACGGGTTAAATATCTGTTATCATCTTTTCAAGTTGGATTTTGGCTCCT 480
 QY 516 TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAAATATGTATGTATATCA 575
 Db 481 TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACAAAAATATGTATGTATATCA 540
 QY 576 AAT 579
 Db 541 AAT 544

RESULT 15
 ACN44176
 ID ACN44176 standard; DNA; 62231 BP.
 ACN44176;
 DT 18-NOV-2004 (first entry)
 XX Mouse genomic sequence MCG15520.
 DE
 XX Cytoskeletal; carcinoma; lymphoma; cancer; murine; gene; ss.
 XX Mus musculus.
 OS
 XX MO2003073826-A2.
 PN 12-SEP-2003.
 PD
 XX 28-FEB-2003; 2003WO-US006235.
 PF
 XX 01-MAR-2002; 2002US-00087192.
 PR
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 PI Morris DW;
 XX WPI; 2003-328604/31.
 DR
 XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.
 PS Claim 1; SEQ ID NO 493; Opp; English.
 XX
 XX The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
 CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
 CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
 CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
 CC determining Carcinoma Associated (CA) gene copy number. In addition, the
 CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
 CC carcinoma including lymphoma. The present sequence is one such CA coding
 CC sequence. Note: This patent is an equivalent to basic Patent
 CC US2002182586A1, for which no sequence data was published
 XX
 SO Sequence 62231 BP; 15482 A; 12591 C; 12729 G; 16514 T; 0 U; 4915 Other;
 Query Match 74.2%; Score 446; DB 11; Length 62231;
 Best Local Similarity 86.1%; Pred. No. 1.2e-120;
 Matches 494; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 28 TCTCCGATGCGCAGATGCTTTTATTCACAAAACCTGCGCGTTACGACCAAAAAGAA 87
 Db 50644 TTCTGACGTGGCTACAGATGCTTTCAATTCAAAACCTGCGCGTTACGACCAAAAAGAA 50703
 QY 88 GATCTTGGGTAATAGGTGTCACAAATCCATCGCCACACCTTAATAGACAGACAGATAG 147
 Db 50704 GATCTTGGGTAATAGGTGTCACAAATCCATCGCCACACCTTAATAGACAGACAGATAG 50763

QY 148 TGAAGTCTGATGATGAGCTTACAGATGACACGAGATACCCCAAAAAGAAAGAGGC 207
 Db 50764 CGAGGTGCTGATGATGAGCTTACAGATGACACGAGATACCCCAAAAAGAAAGAGGC 50823
 QY 208 AGAAGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 267
 Db 50824 GGAAGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 50883
 QY 268 TAATCAGTTTAAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 327
 Db 50884 CAATCAGTTTAAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 50943
 QY 328 GCTTGTCTATGACCGGTGCTAGTTTCAATCAAGTGTGATTAATCTTGAACCGGAATG 387
 Db 50944 GCTTGTCTATGACCGGTGCTAGTTTCAATCAAGTGTGATTAATCTTGAACCGGAATG 51003
 QY 388 ATCCAGGCTGTTAAATGAATGACAGAGATGCTGACCAAAATGATTCAGGCGCACTCAC 447
 Db 51004 GTCCAGGCTGTTAAATGAATGACAGAGATGCTGACCAAAATGATTCAGGCGCACTCAC 51063
 QY 448 TGCCAGTACATGACGAGGTTAATATGCTTTGATCAATTTTTCAGATTGTAATTTT 507
 Db 51064 CGCCAAATGCTCAGAGACGGGTTAATATGCTTTGATCAATTTTTCAGATTGTAATTTT 51123
 QY 508 GCGTGCCTTGTATATTCCTTTTGGGAATTTTAAACCCCACTTAACAAAACATATGTATG 567
 Db 51124 GCGTGCCTTGTATATTCCTTTTGGGAATTTTAAACCCCACTTAACAAAACATATGTATG 51183
 QY 568 TATCAACAAAATGCTTGAAGAGAAACATATTA 601
 Db 51184 CATCAACAAAATGCTTGAAGAGAAACATATTA 51217

RESULT 16
 ACN44177
 ID ACN44177 standard; cDNA; 2087 BP.
 ACN44177;
 DT 18-NOV-2004 (first entry)
 XX Mouse mRNA sequence MCT17722.
 DE
 XX Cytoskeletal; carcinoma; lymphoma; cancer; murine; gene; ds.
 XX Mus musculus.
 OS
 XX MO2003073826-A2.
 PN 12-SEP-2003.
 PD
 XX 28-FEB-2003; 2003WO-US006235.
 PF
 XX 01-MAR-2002; 2002US-00087192.
 PR
 XX (SAGR-) SAGRES DISCOVERY.
 PA
 PI Morris DW;
 XX WPI; 2003-328604/31.
 DR
 XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
 PT comprises a nucleotide sequence.
 PS Claim 1; SEQ ID NO 494; Opp; English.
 XX
 XX The present invention relates to novel DNA and protein sequences which
 CC are associated with carcinomas. The sequences are useful for: (i) for
 CC screening drug candidates; (ii) for screening of bioactive agent capable
 CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
 CC a bioactive agent capable of modulating the activity of CAP; (iv) for
 CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biopro:
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA gene
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published
XX

XX Sequence 2087 BP; 556 A; 487 C; 508 G; 536 T; 0 U; 0 Other;
SQ

Query Match 73.9%; Score 444.4; DB 11; Length 2087;
Best Local Similarity 86.6%; Pred. No. 6.8e-121;
Matches 490; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCGAAAACCTGGCCGTTGAGGACAAAAGAGATTGG 95
DB 508 TGGCTACAGATGCTTTAATTCGAAAACCTGGCCGTTGAGGACAAAAGAGATTCTGG 567
QY 96 GTAAATGGTGTCAATTCATGCGACCACTTAATAGACACAGAGATGAGAGTGC 155
DB 568 GCAAAATGTTATTCATTCATTCATGCGACCACTTAATAGACACAGAGATGAGAGTGC 627
QY 156 TGGATGAGCTCTACAGAGTGAACGAGAGTACCCGAAAACAGAGAGAGAGAGA 215
DB 628 TGATGAGCTGTACAGAGTGAACGAGAGTACCCGAAAACAGAGAGAGAGAGA 687
QY 216 TCATCAAGAACCTCATCAAGAGTGTATCAAGTGGCCATTTTATAGAAATATCACT 275
DB 688 TCATCAAGAACCTCATCAAGAGTGTATCAAGTGGCCATTTTATAGAAATATCACT 747
QY 276 TTAATCAAGATGAGTACGATGATGATGATGATGATGATGATGATGATGATGAT 335
DB 748 TCATCAAGAACCTCATCAAGAGTGTATCAAGTGGCCATTTTATAGAAATATCACT 807
QY 336 TGAACCTGCTCAATTCATCAAGTGTATGATGATGATGATGATGATGATGATGAT 395
DB 808 TGAACCTGCTCAATTCATCAAGTGTATGATGATGATGATGATGATGATGATGAT 867
QY 396 TGTAAATGAATGACAGAGTGTATGATGATGATGATGATGATGATGATGATGAT 455
DB 868 TGTAAATGAATGACAGAGTGTATGATGATGATGATGATGATGATGATGATGAT 927
QY 456 CACATGAGAGGAGTGTATGATGATGATGATGATGATGATGATGATGATGAT 515
DB 928 CTCAGAGAGGAGTGTATGATGATGATGATGATGATGATGATGATGATGAT 987
QY 516 TGTAAATGCTTTTGGGAAATTTTAAACCCCACTTACAAAACCTATGATGATGAT 575
DB 988 TGTAAATGCTTTTGGGAAATTTTAAACCCCACTTACAAAACCTATGATGATGAT 1047
QY 576 AAATGTTGATGAAGAAACATATGA 601
DB 1048 AAATGTTGATGAAGAAACATATGA 1073

RESULT 17
AAH07641
ID AAH07641 standard; cDNA; 587 BP.
XX AAH07641;
XX
XX
XX 26-JUN-2001 (first entry)
XX
XX Human cDNA clone (5'-primer) SEQ ID NO:4476.
XX
XX Human; primer; detection; diagnosis; antitumor therapy; gene therapy; ss.
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX
XX 07-FEB-2001.
PD

XX
PF 28-JUL-2000; 2000EP-00116126.
XX
XX 29-JUL-1999; 99JP-00248036.
XX
XX 27-AUG-1999; 99JP-00300253.
XX
XX 11-JAN-2000; 2000JP-00118776.
XX
XX 02-MAY-2000; 2000JP-00183767.
XX
XX 09-JUN-2000; 2000JP-00241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI; 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX length cDNAs defined in the specification, and for the detection and/or
XX diagnosis of the abnormality of the proteins encoded by the full-length
XX cDNAs.
XX
XX Claim 1, SEQ ID NO 4476; 2537bp + Sequence listing; English.
XX
XX The present invention describes primer sets for synthesizing 5602 full-
XX length cDNAs defined in the specification. Where a primer set comprises:
XX (a) an oligo-dT primer and an oligonucleotide complementary to the
XX complementary strand of a polynucleotide which comprises one of the 5602
XX nucleotide sequences defined in the specification, where the
XX oligonucleotide sequence comprises at least 15 nucleotides; or (b) a combination
XX of an oligonucleotide comprising a sequence complementary to the
XX complementary strand of a polynucleotide which comprises a 5'-end
XX sequence and an oligonucleotide comprising a sequence complementary to a
XX polynucleotide which comprises a 3'-end sequence, where the
XX oligonucleotide sequence comprises at least 15 nucleotides and the combination of
XX the 5'-end sequence/3'-end sequence is selected from those defined in the
XX specification. The primer sets can be used in antisense therapy and in
XX gene therapy. The primers are useful for synthesizing polynucleotides,
XX particularly full-length cDNAs. The primers are also useful for the
XX detection and/or diagnosis of the abnormality of the proteins encoded by
XX the full-length cDNAs. The primers allow obtaining of the full-length
XX cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX represent human amino acid sequences; and AAH1629 to AAH1632 represent
XX oligonucleotides, all of which are used in the exemplification of the
XX present invention
XX

SQ Sequence 587 BP; 178 A; 146 C; 140 G; 116 T; 0 U; 7 Other;
SQ

Query Match 65.7%; Score 394.6; DB 4; Length 587;
Best Local Similarity 96.6%; Pred. No. 2.2e-106;
Matches 422; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 1 CTTGAGCGTCCGCGCGCGTGC-CGACTCTCCAGATGCGCACAGATGCTTTAATCCA 59
DB 105 CTTGAGCGTCCGCGCGCGTGC-CGACTCTCCAGATGCGCACAGATGCTTTAATCCA 164
QY 60 AAAACCTGCGCGCTTACAGAGCAAAAGAGATCTTGGTAAATGAGTGCATCAATCATCG 119
DB 165 AAAACCTGCGCGCTTACAGAGCAAAAGAGATCTTGGTAAATGAGTGCATCAATCATCG 224
QY 120 CCACCACTTAATAGACACAGATGATGATGATGATGATGATGATGATGATGATGAT 179
DB 225 CCACCACTTAATAGACACAGATGATGATGATGATGATGATGATGATGATGATGAT 284
QY 180 GGGAGTACACCCAAAACAGAGAGAGAGAGAGATCATCAAGAACTCATCAAGACAG 239
DB 285 GGGAGTACACCCAAAACAGAGAGAGAGAGAGATCATCAAGAACTCATCAAGACAG 344
QY 240 TCATCAAGTGGCCATTTCTTATAGAAATATCAATGATGATGATGATGATGATGAT 299
DB 345 TCATCAAGTGGCCATTTCTTATAGAAATATCAATGATGATGATGATGATGATGAT 404
QY 300 TGGAGAAATTTAAGAAAGTTCATCAGCTGTATGACCGTGTGATTCATCAGG 359

Db 405 TGGAGAAATTTAGAAAGTTGATCAGCTGTGTATGACCGTGGNAGTTCCAAAG 464
 QY 360 TGGATTATACCTTTGACCGGATGTGTTATTCAGGCTGTTAATGAATGCAG-AGAGATG 418
 Db 465 GGGGTTATACCTTTGACCGGATGTGTTATTCAGGCTGTTAATGAATGCANNAAGAAATG 524
 QY 419 CTGACCAATTCATTCGA 435
 Db 525 CTGACCAAAACANTTA 541

RESULT 18

AD063108
 ID AD063108 standard; cDNA; 1986 BP.

AC AD063108;

DT 07-OCT-2004 (first entry)

DE Novel human cDNA sequence #269.

XX ss; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
 KM cytostatic; gene therapy; diagnostic marker; morbid state; osteoporosis;
 KM neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
 KM cancer.

XX Homo sapiens.

XX EP1440981-A2.

XX 28-JUL-2004.

XX 21-JAN-2004; 2004EP-00001196.

XX 21-JAN-2003; 2003JP-00102206.

PR 09-MAY-2003; 2003JP-00131392.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Nagai K, Irie R;

XX WPI; 2004-535376/52.

DR P-PSDB; AD065236.

PT Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.

PS Claim 1; SEQ ID NO 269; 2449bp; English.

XX The invention relates to 2495 novel polynucleotides (I) and their encoded
 CC polypeptides, sequences hybridizing to these nucleotides, sequences
 CC encoding partial polypeptides and sequences having 70% or 90% identity to
 CC the nucleotide and protein sequences. The nucleotides and polypeptides
 CC are useful as diagnostic markers or therapeutic target for the diseases
 CC or morbid states. They are also useful for treating osteoporosis,
 CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,
 CC dementia and various cancers. This sequence corresponds to a nucleotide
 CC sequence of the invention.

XX Sequence 1986 BP; 536 A; 458 C; 497 G; 495 T; 0 U; 0 Other;

Query Match 42.3%; Score 254.4; DB 12; Length 1986;

Best Local Similarity 65.9%; Pred. No. 1.6e-64;

Matches 369; Conservative 0; Mismatches 191; Indels 0; Gaps 0;

QY 42 CAGATGCTTTATTCAGAAACCTGGCCCTTACGACACAAAGAGATTTGGTAAAA 101

Db 150 CTGATGCTTTATTCAGAAAGCTTGGCGCTTCAAGCCGAGAGAGATTCGAGCAAA 209

QY 102 TGGTGTCAAATCATGCGACACCTTATAGACGACAAAGTAGAGGTCTGATG 161

Db 210 TAGCAGCAAAAGCTGTGGCCACATGTGATTGATGACACAGACGAGATCTTTGATG 269
 QY 162 AGCTTACAGAGTGACCGAGAGTACACCCAAACAAAGAGGACGAGAGATCTATCA 221
 Db 270 AGCTTACAAAGTACACCAAGAGACACACACAAAGAGAGCCACAGATCTATGA 329
 QY 222 AGAACCTCATCAGACAGTATCAGCTGGCCATTCTTTAGGAATTAATCAGTTTATG 281
 Db 330 AAGACTTAATCAAGTGGCGATCAAAATCGGATCTCTACCGAACACACAGTTTATGCC 389
 QY 282 AAGATGAGCTAGCATGATGAGAGAAATTTAAGAGAAAGTTTACAGCTTGTATGACCG 341
 Db 390 AAGAGAGAGCTGTTATTTGTGAGAGAGTCCGGAAGAAAGCTGAAACCGCATGACCA 449
 QY 342 TGGTCAAGTTTCAATCAGGTGATTTATCCTTTACCGGAATGTGTTATCCAGGCTGTAA 401
 Db 450 TTGTCAAGCTTCTATGAGGTGATATACCTTCATGAGAAAGTGTCTCCATCTCTGC 509
 QY 402 ATGAATGCAGAGAGAGTGTGACCAACCAATCATTTAGGCGCCCTCAGTCCCAAGTCATG 461
 Db 510 ATGATGCAGAGAGCTGTGTGATGAATCTGTGACGCGCACCTGACGCCACGACGACCG 569
 QY 462 GACGGGTTAATATGCTTTGATCATTTTTCAGATTTGATTTTGGCTGCTGTATTA 521
 Db 570 GGGGCTACCAACCGTCTTTTACCACTTGTCCGATGTGAGATTCTCTCCACCTCTATA 629
 QY 522 ATCCTTTGGGAATTTTAAACCCCACTTACAAAAACCTATGTGATGATCAACAAATGT 581
 Db 630 GTCTGATGAGAGCTGTAGGCCCAACCTCAAGAGGATTTGTGAAGATCATATATAGTTC 689
 QY 582 TGGATGAAGAGAAATATGA 601
 Db 690 TAGATGAGAAAGTCTTTTAA 709

RESULT 19

AD030769
 ID AD030769 standard; cDNA; 969 BP.

XX AD030769;

DT 18-DEC-2003 (first entry)

DE Human novel cDNA sequence, SEQ ID NO:851.

XX Human; diagnostic; drug screening; forensics; gene mapping;
 KM biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KM neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KM ulcers; osteoporosis; autoimmune disease; cancer;
 KM molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KM neuroprotective; anti-nausea; anticoagulant; thrombolytic; vulnerary;
 KM anticancer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KM gene therapy; chromosome 15q21.2; gene; ss.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSR-) HYSRQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;
 PI Haley-Vicente D, Drmanac RT;

XX WPI; 2003-371981/35.

DR P-PSDB; AD031740.

XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.

XX Claim 1; SEQ ID NO 851; 1185bp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 98% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridization probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 969 BP; 264 A; 261 C; 257 G; 187 T; 0 U; 0 Other;

Query Match 41.6%; Score 249.8; DB 10; Length 969;
 Best Local Similarity 66.4%; Pred. No. 2.6e-63;
 Matches 359; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 42 CAGATGCTTTTAATTCAAAACCTGGCCGTTGAGCACAAGAAAGATCTGGTAA 101
 DB 320 CTGATGTTTTTGTTCAAAAGAGCTTGCCTTCAGCCCAAGAAAGATCTGACAAA 379
 QY 102 TGGTGTCCAAATCCATCGCACCACTTAATGAGACAACAAGTATGAGGTGCTGATG 161
 DB 380 TAGCCAGCAAAATGTGGCCCAATGTGATATATACACGACACGAGATCTTGAAGT 439
 QY 162 AGCTTACAGATGACCCAGGAGTACACCCAAAACAAGAGAGGAGAGATCATCA 221
 DB 440 AGCTTACAAAAGTCAACCAAGAGCAACACAAAGAGAGAGCCCAATATATCA 499
 QY 222 AGAATCTATCAAGACATCATAGCTGGCCATTTTATAGAAATTAATCACTTTAATC 281
 DB 500 AAGACTTAATCAAGGTGGGATCAAAATGGGATCCTCTACCGAAACAACAGTTTATGACC 559
 QY 282 AAGATGAGTATGATGAGAAATTTAAGAAAGAAAGTCACTACGTTGCTATAGACG 341
 DB 560 AAGAGAGCTGTGTTATTTGAGAAAGTCCGGAAGAGCTGAACCAAGCCATACCA 619
 QY 342 TGGTCAATTCATCAGTGTGATTAATCCTTTGACCGGATATGTTATCCAGGCTTTAA 401
 DB 620 TTGTGAGCTTTATAGGTGAATACACCTTGATAGGAACGTGCTCCAAATCTCTGC 679
 QY 402 ATGAATGACAGAGATGCTGCACCAATATCAATTTAGGGCCCACTCACTGCCAATCATAGT 461
 DB 680 ATGAGTGAAGGACCTGTGTCATGAATCTGGTGCAGGCGACCTGAGCCGACAGACCACG 739

QY 462 GACGGTTAATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGATA 521
 DB 740 GCGCATCAACCAAGCTCTTTAACACTTTGCCGATGTGAGTCTCTCCACCTCTATA 799

QY 522 ATCTTTTGGGAAATTTTAAACCCCACTTACAAAACATATGTGATGATCAACAAATGT 581
 DB 800 GTCTGATGAGACTGTAGGCCCACTCAAGAGATTTGTGAAGAAATCAATATAGTTGC 859

QY 582 T 582
 DB 860 T 860

RESULT 20

ACH92216
 ID ACH92216 standard; DNA, 527 BP.

ACH92216;

29-JUL-2004 (first entry)

Human genome derived single exon probe #25411.

Human; probe; ss; gene expression; single exon probe; microarray;
 alternative splicing event; genomic alteration.

Homo sapiens.

US2003194704-A1.

16-OCT-2003.

03-APR-2002; 2002US-00029386.

03-APR-2002; 2002US-00029386.

(PENN/) PENN S G.

(RANK/) RANK D R.

(HANZ/) HANZEL D K.

Penn SG, Rank DR, Hanzel DK;

WPI; 2004-119264/12.

New human genome-derived single exon nucleic acid probes useful for human
 gene expression analysis, for identifying or characterizing alternative
 splicing events, for assessing genomic alterations or as tools for
 surveying tissues.

Claim 1; SEQ ID NO 25411; 80bp; English.

The invention relates to a nucleic acid probe for measuring human gene
 expression, comprising any of the 27,400 fully defined nucleotide
 sequences in the specification, or their complements or fragments, and
 encoding at least 8 amino acids of any of the 688 amino acid sequences
 fully defined in the specification. The probe is a single exon probe that
 hybridizes under high stringency conditions to a nucleic acid molecule
 expressed in human cells or tissues. Also included are a spatially-
 addressed set of single exon nucleic acid probes for measuring human
 gene expression (comprising a plurality of single exon nucleic acid
 probes cited above, where each of the plurality of probes is separately
 and addressably isolatable or amplifiable from the plurality), a single
 exon microarray for measuring human gene expression, a method of
 measuring human gene expression, a vector comprising the single exon
 probe cited above, an ORF-encoded peptide comprising at least 8
 contiguous amino acids of any of the above-mentioned amino acid
 sequences (optionally with conservative amino acid substitutions), an
 isolated antibody that binds specifically to a peptide cited above,
 methods of selling and/or licensing single exon probes or microarrays to
 a customer desiring to measure gene expression, a method of providing
 human gene expression data by subscription, and a computer-readable
 storage medium which contains a database having a plurality of records

CC (each record including data on the expression of a single exon probe
CC cited above. The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the invention. Note: The sequence data for this
CC in electronic format directly from USPRO at
CC seqdata.uspro.gov/sequence.html?docID=20030194704
XX
SQ Sequence 527 BP; 163 A; 129 C; 128 G; 107 T; 0 U; 0 Other;
Query Match 39.5%; Score 237.4; DB 12; Length 527;
Best Local Similarity 65.7%; Pred. No. 9.1e-60;
Matches 346; Conservative 0; Mismatches 181; Indels 0; Gaps 0;
QY 75 AGGCACAAAAGAAATCTTGCTTAAATGCTTCAATCCATCCGCCACCTTATAG 134
1 AAGCCAGAAAGAAATCTTGAGCAAAATATCCAGCAAACTGTGGCCACATGTTGATTG 60
QY 135 ACGACACAAATAGTAGGCTGCTGGATGAGCTTACAGAGTGACCGAGGAGTACACCCAA 194
61 ATATACACACACAGCATCTTGTATGATGAGCTTCAAAAGTACCACCAAGACACACACA 120
QY 195 ACAAGAAGAGGAGAGAAATCATCAAGAACTCATCAAGACAGCTATCAAGCTGGCCA 254
121 ACAAGAAGAGAGCCCAAGATCATGAAAGACTTAATCAAGTGGCGATCAAAATGCGGA 180
QY 255 TTCTTATAGAAATTAATCAAGTTAATCAAGATGAGTATGATGAGAGAAATTTAGA 314
181 TCCTTACCGGAACAACATTTAGCCAAAGAGAGCTGTTATTTGGAAAGTCCGGA 240
QY 315 AGAAAGTTATCACTGACTTGTATATACCGTGTCACTTTCATCAAGTGGATTATACCTTGG 374
241 AGAAGCTGAACCAAGACCGCATATGCTTCACTTCTATGAGTGAATATACCTTGG 300
QY 375 ACCGGAATGTTATTCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCATTC 434
301 ATAGGAACGTCCTCTCAATCTCTGATGATGCAAGAGCTGTGATCAAGAACTGGTGC 350
QY 435 AGCGCACTCTGACCAATGACATGACGAGGTTATATATGCTTTGATCATTTTTCAG 494
361 AGCGGACCTGAGCGCCAGAACCCAGCGGCATCAACACGCTTTTACCACTTTGCCG 420
QY 495 ATTGTGAATTTTGGCTGCTGCTGTATATCTTTTGGGAATTTTAAACCCACTTACAA 554
421 ATGTGAGTTCCTCTCCACCTCTATATGCTGATGAGAGCTGTAGGCCCAACCTCAGA 480
QY 555 AACTATGATGATGATCAAAATGTTGATGAAGAAACATATGA 601
481 GGATTTGTGAAGAAATCAATGAATGCTATAGATGAAGAAAGTCTTTAA 527
DB
RESULT 21
ABS78741
ID ABS78741 standard; DNA; 645 BP.
XX
XX ABS78741;
AC
XX
XX
DT 16-DEC-2002 (first entry)
XX
XX
DE DNA encoding human NOVX17a protein.
XX
XX Human; NOVX; human disease; NOVX-associated disorder; cancer; addiction;
KW Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
KW tuberculous sclerosis; hypercalcaemia; Parkinson's disease; depression;
KW Huntington's disease; cerebral palsy; epilepsy; Lesch-Nyhan syndrome;
KW multiple sclerosis; ataxia-telangiectasia; leukodystrophy; anxiety; pain;

KW obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
KW infertility; inflammatory bowel disease; atherosclerosis; hypertension;
KW scleroderma; hemophilia; diabetes; pancreatitis; autoimmune disease;
KW asthma; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
KW bacterial infection; parasitic infection; graft-versus-host disease;
KW cell differentiation; cell proliferation; haematopoiesis; wound healing;
KW angiogenesis; gene; de.
XX
XX Homo sapiens.
OS
XX
PN WO200272770-A2.
XX
PD 19-SEP-2002.
XX
XX
PF 08-MAR-2002; 2002WO-US007283.
XX
XX
PR 08-MAR-2001; 2001US-0274281P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 20-MAR-2001; 2001US-0277239P.
PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277919P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0279366P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-MAY-2001; 2001US-0288148P.
PR 31-MAY-2001; 2001US-0294821P.
PR 31-OCT-2001; 2001US-0335302P.
PR 04-DEC-2001; 2001US-0338375P.
PR 07-MAR-2002; 2002US-00094466.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Spytek KA, Vernet CA, Tchermey VT, Malyankar UM, Gerlach VL;
PI Li L, Zernusen BD, Paturajan M, Gusev VY, Kekuda R, Pena CBA;
PI Zhong M, Gangoli EA, Taupier RJ;
XX
XX WPI; 2002-713508/77.
XX P-PSDB; ABG97497.
XX
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders, e.g. diabetes, multiple
PT sclerosis, atherosclerosis, cancer, infections, osteoporosis or
PT Parkinson's disease.
XX
XX
PS Claim 22; Page 139; 266pp; English.
XX
XX The present invention relates to a new polypeptide (NOVX). The NOVX
CC polypeptide, nucleic acid and antibody are useful in the manufacture of a
CC medicament for treating a syndrome associated with a human disease,
CC preferably a NOVX-associated disorder. The NOVX nucleic acids,
CC polypeptides and antibodies are useful for treating, preventing or
CC diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
CC syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-
CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
CC infertility, inflammatory bowel disease, atherosclerosis, hypertension,
CC scleroderma, haemophilia, diabetes, pancreatitis, autoimmune disease,
CC asthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
CC infections, or graft-versus-host disease. The nucleic acids and
CC polypeptides may also be used as targets for the identification of small
CC molecules that modulate or inhibit e.g. neurogenesis, cell
CC differentiation, cell proliferation, haematopoiesis, wound healing and

CC angiogenesis, in gene therapy, in generation of antibodies that bind
 CC immunospecifically to NOXV substances for use in therapeutic or
 CC diagnostic methods. The nucleic acids are further used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine, and
 CC pharmacogenomics. The present nucleic acid sequence encodes a human NOXV
 CC protein of the invention

XX Sequence 645 BP; 97 A; 228 C; 219 G; 101 T; 0 U; 0 Other;

Query Match 34.6%; Score 207.8; DB 6; Length 645;
 Best Local Similarity 59.9%; Pred. No. 6.1e-51;
 Matches 347; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 23 CCGACTCTCCGATGCGCACAGATGCTTTAATTCAAAAAAGCTGGCCGTTGAGGACAA 82
 DB 8 CTGTGTCCCCCGAGGCGCATGACACCTTCAGACCAAGAGCTGCTGCGAGGGCGAG 67
 QY 83 AAGAGATCTTGGGTAAATGTTGTCGAATTCATGCGCACCACTTAAATAGACACACA 142
 DB 68 AAGAGATCTCTGTAGTAAGATGCGCTCAAGGCAAGTGGCGCTGCTGGTGAATGACACC 127
 QY 143 AGTAGAGAGTGTGATGATGAGCTCTACAGATGACAGGAGAGTACACCAAAACAGAAAG 202
 DB 128 AGCAGTAGAGTGTGATGATGAGCTCTACAGATGACAGGAGAGTACACCGGACCGCAAG 187
 QY 203 GAGGACAGAGATCATCAAGAACTCATCAAGACAGTCAATCAAGCTGCGCATTTTAT 262
 DB 188 GAGGCCCAAGAAATGCTCAAGAACTGCTCAAGAGTGGCCCTGTAAGCTGGAGATGCTGCTG 247
 QY 263 AGGAATTAATCATGTTAATCAAGATGAGTACAGTGAAGAAATTTAAGAAAGAAAGTT 322
 DB 248 CGTGGGACACAGCTGGGCGGTGAGAGCTGGCGCTGCTGGCGGCTTCGCGCACCGGCGG 307
 QY 323 CATCACTGTATGACCGTGTGATTTCCATCAAGTGGATTAATCTTTGACCGGAAT 382
 DB 308 CGTGTGCTGCAATGACGCGCTGACCTTCACAGGTGACTTCACTTCGACCGGCGC 367
 QY 383 GTGTTATCCAGGCTGTTAATGATGACAGAGATCTGCGACCAAAATCATTCAGCGCCAC 442
 DB 368 GTGCTGCGCGCGGGTGTCTGAGTCCCGGACCTCTGACCAAGCCGCGGGTCCCGAC 427
 QY 443 CTCACCTCCAAAGTCAATGACGAGGCTTAATATGCTTTGATCATTTTTCAGATTGAA 502
 DB 428 CTGACCGCAAGTCCACCGGCGCATCAACCAAGTTCGGCGACCTAGCGGACCTCGAC 487
 QY 503 TTTTGGCTGCTGTATATATCTTTTGGGAATTTAAACCCACTTACAAAACATATG 562
 DB 488 TTCTGTGGCTGCTGTATACGCGCCCGCGGACCTTACCGCTCCACCTGCGGAGATCTGC 547
 QY 563 GATGTATCAACAAATGTTGATGAAGAGACATATGA 601
 DB 548 GAGGGCTGGGCGGATGCTGACGAGGCGAGCGCTCTGA 586

RESULT 22

AA139626
 ID AA139626 standard; DNA; 2108 BP.

XX AA139626;

05-SEP-2002 (first entry)

XX Human secreted protein DNA SEQ ID NO 70.

DE Antiarteriosclerotic; cytoskeletal; HIV; antiallergic; antianemic;
 XX antiasthmatic; cardiast; vasotropic; neuroprotective; nootropic; SECP;
 KW anticonvulsant; antiparkinsonian; cerebroprotective; antiinflammatory;
 KW immunosuppressive; human secreted protein; cell proliferative disorder;
 KW arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;
 KW allergy; anaemia; asthma; cardiovascular disease; developmental disorder;
 KW ischaemic heart disease; congestive heart failure; neurological disorder;
 KW renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;
 KW Parkinson's disease; epilepsy; stroke; knockin humanised animal;

KW transgenic animal; gene therapy; gene; ds.

XX Homo sapiens.

OS WO200238602-A2.

XX 16-MAY-2002.

PD 08-NOV-2001; 2001WO-US047420.

PF 08-NOV-2000; 2000US-0247505P.

PR 09-NOV-2000; 2000US-0248642P.

PR 16-NOV-2000; 2000US-0249824P.

PR 21-NOV-2000; 2000US-0252824P.

PR 08-DEC-2000; 2000US-0254305P.

PR 18-DEC-2000; 2000US-0256448P.

XX (INCY-) INCYTE GENOMICS INC.

PA Yue H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Wallia NK;
 PI Sanjanwala M, Thornton M, Elliott VS, Lu Y, Gietzen KU, Burford N;
 PI Ding L, Hafalia AJA, Tang YT, Bandman O, Warren BA, Honchell CD;
 PI Lu DM, Thangavelu K, Lee S, Xu Y, Yang J, Lal PG, Tran B;
 PI Ison CH, Duggan BM, Saperstein SK;
 XX MPI; 2002-519296/55.
 DR P-PSDB; AAO21665.

XX Human secreted proteins and polynucleotides for diagnosing, treating or
 PT preventing disorders of cell proliferative, cardiovascular,
 PT developmental, neurological and autoimmune/inflammatory disorders.

PS Claim 5; Page 195-196; 22pp; English.

XX The invention relates to an isolated human secreted protein (SECP)
 CC polypeptide from 63 fully defined protein sequences given in the
 CC specification. The polypeptide is useful for the diagnosing/treating of a
 CC disease with decreased/overexpression of SECP. Examples of disorders
 CC associated with abnormal expression of SECP include a cell proliferative
 CC disorder e.g. arteriosclerosis, cancers; autoimmune/inflammatory
 CC disorder, AIDS, allergies, anaemia, asthma; cardiovascular disease e.g.
 CC congestive heart failure, ischaemic heart disease; developmental disorder
 CC e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.
 CC Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.
 CC The SECP polynucleotide and polypeptide are further useful for analysing
 CC the proteome of a tissue or a cell type. The polynucleotide is useful for
 CC creating knockin humanised animals (pigs) or transgenic animals (mice or
 CC rats) to model human disease, and for somatic or germline gene therapy,
 CC and further for generating hybridisation probes useful in mapping the
 CC naturally occurring genomic sequence. This polynucleotide sequence
 CC represents the DNA of a human secreted protein of the invention

XX Sequence 2108 BP; 408 A; 602 C; 636 G; 462 T; 0 U; 0 Other;

QY Query Match 34.6%; Score 207.8; DB 6; Length 2108;
 Best Local Similarity 59.9%; Pred. No. 1.1e-50;
 Matches 347; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 23 CCGACTCTCCGATGCGCACAGATGCTTTAATTCAAAAAAGCTGGCCGTTGAGGACAA 82
 DB 412 CTGTGTCCCCCGAGGCGCATGACACCTTCACACCAAGAGCTGCTGAGGCGGCGAG 471
 QY 83 AAGAGATCTTGGGTAAATGTTGTCGAATTCATGCGCACCACTTAAATAGACACACA 142
 DB 472 AAGAGATCTCTAGTAAGTGGCGTCAAGGAGTGGCCGCTGCTGGATGATGACACC 531
 QY 143 AGTAGAGAGTGTGATGATGAGCTCTACAGATGACAGGAGAGTACACCAAAACAGAAAG 202
 DB 532 AACAGTAGAGTGTGATGATGAGCTCTACAGATGACAGGAGAGTACACCGGACCGGCAAG 591
 QY 203 GAGGACAGAGATCATCAAGAACTCATCAAGACAGTCAATCAAGCTGCGCATTTCTTAT 262
 DB 592 GAGGCCCAAGAAATGCTCAAGAACTGCTCAAGAGTGGCCCTGAAGCTGGAGTGGTCTG 651

QY 263 AGGAATAATGATTTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGTT 322
 DB 652 CTTGGGAGCAAGCTGGCGCGGTGAGAGCTGCGTCTGGCGCGCTTCCGCAACCGGCG 711
 QY 323 CATCAGCTTCTGATGACCGTGTGATTTCCATCAGGTGATTTATCTTTGACCGGAAT 382
 DB 712 CGCTGCGCTGCGCATGAGCGCGCTGAGCTTCCACAGGTGATTTACCTTCCACCGGCG 771
 QY 383 GTGTTATTCAGGCTGTTAAATGAATGACAGAGATGCTGACCAAAATTCATTCAGCGCAC 442
 DB 772 GTCTGCGCGCGCGGCTGCTGAGTGCAGGACCTGCTGACAGCGCGGTGGTCCCGCAC 831
 QY 443 CTGACGCGCAAGTCACATGACCGGTTAATATGCTTTTATCATTTTTCAGATTGGA 502
 DB 832 CTGACCGCGCAAGTCACAGCGCGCGATCAACCATGTTGCGCACCTGACGCTGCGAC 891
 QY 503 TTTTGGCTGCTGTTTATTAATCTTTTGGGAATTTTAAACCCCACTTACAAAACTATGT 562
 DB 892 TTCTGCTGCTGCTGCTTACGCGCGCGCGGACCCCTTACCGCTTCCACCTGCGCAGATCTGC 951
 QY 563 GATGTATCAACAATAATGTTGATGAGAGAAATATGA 601
 DB 952 GAGGCGCTGGGCGGATGCTGACGAGGAGGAGCGCTTGA 990

RESULT 23
 ADQ85659
 ID ADQ85659 standard; cDNA; 2186 BP.
 AC ADQ85659;
 DT 07-OCT-2004 (first entry)
 DE Human tumour-associated antigenic target (TAT) cDNA sequence #2473.
 XX human: tumour-associated antigenic target; TAT; cytostatic; gene therapy;
 KM cancer; cell proliferative disorder; gene; ss.
 OS Homo sapiens.
 XX MO2004060270-A2.
 PN 22-JUL-2004.
 XX 15-OCT-2003; 2003WO-US029126.
 PF 18-OCT-2002; 2002US-0418988P.
 PR (GENTH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 PI Wu TD, Zhou Y;
 XX MPI, 2004-534300/51.
 DR New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 PS Claim 1; SEQ ID NO 2473; 5504bp; English.
 XX The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO.1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence

CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.

SQ Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;
 Query Match 34.0%; Score 204.2; DB 12; Length 2186;
 Best Local Similarity 59.3%; Pred. No. 1.3e-49;
 Matches 347; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 17 CGGTGCGGACCTCTCGATGCGCACAGATGCTTTAATTTCAAAACCTGCGCTTGC 76
 DB 92 CGACTCTGCTTGAAGATGAGGATGACATCTTACGACCAAGACCTGCGCTTGCAG 151
 QY 77 GCACAAAGAGATCTTGGGTAATATGCTGCCAATTCATCGCCACCACTTAATATGAC 136
 DB 152 GCGCAAGAGAGCTCTGATGATGAGATGGGTCCAAAGCGATGGTGGCGTGGTGGAT 211
 QY 137 GACACAAGTATGAGGTGCTGATGAGCTTACAGATGACAGGAGATACCCAAAC 196
 DB 212 GACACCAGAGATGAGGTGCTGATGAGCTGTACCGGCCACCGAGATTCACGCCAGC 271
 QY 197 AAGAAGAGAGAGAGATGATCATCAAGAACCTCATGAAGACATCAAGCTGGGCATT 256
 DB 272 CGCAAGAGAGAGAGAGATGCTCAAGAACCTGTGAAGGTGCTTGAAGCTGGAGACTG 331
 QY 257 CTTATAGAGATTAATGATTTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAG 316
 DB 332 CTGCTGCGTGGGAGACAGCTGCGGCGGTGAGAGCTGGCCGTGCGCGCTTCCGCCAC 391
 QY 317 AAGTTCAATCACTGCTGATGACCGTGTGATTTTCAATCAGGTGATTAATCACTTTGAC 376
 DB 392 CGGCGCGCTGCTGCTGATGACGCGGCTGACGCTTCAAGGTGAGCTTCACTTGCAC 451
 QY 377 CGGAATGCTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGACCAAAATTCATTCAG 436
 DB 452 CGGCGGTGCTGAGCGCGCGGCTGCTGAGTCCCGGACCTGTGACCAAGGCGCGTGGT 511
 QY 437 CGCAGCTACTGCGCAAGTCACATGAGCGGTTAATATATGCTTTTATCATTTTTCAGAT 496
 DB 512 CCGCAGCTACCGCGCAAGTCCACGCGCGCATCAACAGATGCTGGCGCACTTACGCGAC 571
 QY 497 TGTGAATTTTGGCTGCTGCTGATTAATCTTTTGGAAATTTTAAACCCCATTTCAAAA 556
 DB 572 TCGGACTTCTGCTGCTGCTTACGCGCGCGCGGACCTTACCGCTCCACCTGCGCAGG 631
 QY 557 CTATGATGATGATTAACAATAATGTTGATGAGAGAGATATGA 601
 DB 632 ATCTGCGAGGCGCTGGGCGGATGCTGAGCGAGGAGGAGCTTGA 676

RESULT 24

AD086751 standard; cDNA; 2186 BP.

AD086751;
07-OCT-2004 (first entry)

Human tumour-associated antigenic target (TAT) cDNA sequence #3626.

human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
cancer; cell proliferative disorder; gene; ss.

Homo sapiens.

WO2004060270-A2.

22-JUL-2004.

15-OCT-2003; 2003WO-US029126.

18-OCT-2002; 2002US-0418988P.

(GENENTECH) GENENTECH INC.

(WU) WU T D.

(ZHOU) ZHOU Y.

Wu TD, Zhou Y;

WPI; 2004-534300/51.

New nucleic acid molecule and encoded polypeptide, for diagnosing,

preventing or treating cell proliferative disorders such as cancer.

Claim 1; SEQ ID NO 3626; 5504bp; English.

The present invention describes an isolated tumour-associated antigenic target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of (a); (c) the complement of (a) or (b); (d) a sequence that has 80% sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-(c). Also described: (1) an expression vector comprising the above nucleic acid; (2) a host cell comprising the above expression vector; (3) a process for producing a polypeptide; (4) an isolated polypeptide comprising: (a) an amino acid sequence encoded by any of the above nucleotide sequences; (b) an amino acid sequence encoded by the full-length coding region of the above nucleotide sequences; or (c) a sequence having at least 80% identical to (a) or (b); (5) a chimeric polypeptide comprising the above polypeptide fused to a heterologous polypeptide; (6) an isolated antibody that binds to the above polypeptide; (7) a process for producing the antibody; (8) an isolated oligopeptide that binds to the above polypeptide; (9) a tumour-associated antigenic target (TAT) binding organic molecule that binds to the above polypeptide; (10) a composition of matter comprising the above (chimeric) polypeptide, antibody, oligopeptide or TAT binding organic molecule, in combination with a carrier; (11) an article of manufacture comprising a container and the composition of matter contained within the container; (12) methods of inhibiting the growth of a cell that expresses the above protein, where the growth of the cell is at least in part dependent upon a growth potentiating effect of the above protein; (13) a method of therapeutically treating a mammal having a cancerous tumour comprising cells that express the above protein; (14) a method of determining the presence of a protein in a sample suspected of containing the protein described above; (15) methods of diagnosing the presence of a tumour in a mammal; (16) a method for treating or preventing a cell proliferative disorder associated with increased expression or activity of the above protein; and (17) a method of binding an antibody, oligopeptide or organic molecule to a cell that expresses the protein described above. The TAT sequences have cytostatic activities, and can be used in gene therapy. The composition and methods are useful for diagnosing, preventing or treating cancer. The composition is also used for preparing

CC a medicament for the therapeutic treatment or diagnostic detection of a
CC cell proliferative disorder or cancer. The present sequence represents a
CC human TAT cDNA sequence from the present invention.

SQ Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;

Query Match 34.0%; Score 204.2; DB 12; Length 2186;
Best Local Similarity 59.3%; Pred. No. 1.3e-49;
Matches 347; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 17 CCGTCGCGGACTCTCCGATGGCCACAGATGCTTTAAATTCAAAACCTGCGCTTACG 76
DB 92 CGAGCTGCTTCGAGATGAGCCATGACACTTCACACCAAGAGCTGCTGTCAG 151
QY 77 GCACAAAGAGATCTTGCTAAATGATGCCAATTCATGCCACCACTTAATAGAC 136
DB 152 GCGCAAGAAAGCTCTGATGAATGAGCTGCAAGGAGATGAGCGCTGCTGATGAT 211
QY 137 GACACAAAGTATGAGGTGCTGATGAGCTTCACAGATGACCGAGATACACCAAAAC 196
DB 212 GACACCAAGATGAGGTGCTGATGAGCTTCACCGGACCAAGGAGTTCACCGGACG 271
QY 197 AAGAGAGGAGCAGAGAGATCATCAAGAACTCATCAAGACAGTATCAAGCTGCGCAT 256
DB 272 CGCAAGGAGGCCACAGAGATGCTCAAGAACTGCTCAAGTGGCCCTGAGAGCTGGAGCTG 331
QY 257 CTTTATAGGAATTAATCAGTTTAATCAGATGAGTGAATGAGAGAAATTAAGAG 316
DB 332 CTGCTGCGGAGGACAGCTGGGCGGTGAGAGCTGAGCTGCTGCGGCTTCCGCGAC 391
QY 317 AAGATTCATCAGCTTGTATGACCGTGGTCACTTTCATCAGTGGATTAATACCTTTCAG 376
DB 392 GGGGCGCGCTGCTGCGCATGACGCGCTGACCTTCACCAAGTGAATTCACCTTCGAC 451
QY 377 CGGAATGTGTTATCCAGAGCTGTTAAATGATGACAGAGATCTGCACTCAATTCAG 436
DB 452 CGCGCGTGTGCGCGCGCGCTGCTGAGTGCCTGCTGACCTGCTGACCGAGCGCGTGG 511
QY 437 GCGCACTCAGCTGCGCAAGTCAATGACGCGGTAAATGCTTGAATTTTACGAT 496
DB 512 CCCACCTGACCGCAAGTCCACGCGCATCAACAGTGTGGCTACCTAGCGAC 571
QY 497 TGTGAATTTTGGCTGCTGCTGATATATCTTTTGGGAATTTTAAACCCCACTTACAAA 556
DB 572 TCGACTTCCTGCTGCTGCTGCTGATAGGCGCCCGGAGCCCTTCCGCTCCACCTGGCAGG 631
QY 557 CTATGTGATGTATCAACAAATGTTGATGAAGAGAAATATGA 601
DB 632 ATCTCGAGGAGCTGCGCGATGCTGACGAGGCGACGCTCTGA 676
RESULT 25
AD084271
ID AD084271 standard; cDNA; 2186 BP.
AC AD084271;
DT 07-OCT-2004 (first entry)
DE Human tumour-associated antigenic target (TAT) cDNA sequence #1085.
KW human; tumour-associated antigenic target; TAT; cytostatic; gene therapy;
cancer; cell proliferative disorder; gene; ss.
OS Homo sapiens.
PN WO2004060270-A2.
PD 22-JUL-2004.
PF 15-OCT-2003; 2003WO-US029126.
PR 18-OCT-2002; 2002US-0418988P.

XX (GETH) GENENTECH INC.
 PA (WUTD/) WU T D.
 PA (ZHOU/) ZHOU Y.
 PI WU TD, Zhou Y;
 XX WPI, 2004-534300/51.
 DR
 XX New nucleic acid molecule and encoded polypeptide, for diagnosing,
 PT preventing or treating cell proliferative disorders such as cancer.
 XX
 XX Claim 1; SEQ ID NO 1085; 5504bp; English.
 CC The present invention describes an isolated tumour-associated antigenic
 CC target (TAT) nucleic acid comprising: (a) any of 4622 nucleotide
 CC sequences (see SEQ ID NO:1 to 4622); (b) the full-length coding region of
 CC (a); (c) the complement of (a) or (b); (d) a sequence that has 80%
 CC sequence identity to (a)-(c); or (e) a sequence that hybridises to (a)-
 CC (c). Also described: (1) an expression vector comprising the above
 CC nucleic acid; (2) a host cell comprising the above expression vector; (3)
 CC a process for producing a polypeptide; (4) an isolated polypeptide
 CC comprising: (a) an amino acid sequence encoded by any of the above
 CC nucleotide sequences; (b) an amino acid sequence encoded by the full-
 CC length coding region of the above nucleotide sequences; or (c) a sequence
 CC having at least 80% identical to (a) or (b); (5) a chimeric polypeptide
 CC comprising the above polypeptide fused to a heterologous polypeptide; (6)
 CC an isolated antibody that binds to the above polypeptide; (7) a process
 CC for producing the antibody; (8) an isolated oligopeptide that binds to
 CC the above polypeptide; (9) a tumour-associated antigenic target (TAT)
 CC binding organic molecule that binds to the above polypeptide; (10) a
 CC composition of matter comprising the above (chimeric) polypeptide,
 CC antibody, oligopeptide or TAT binding organic molecule, in combination
 CC with a carrier; (11) an article of manufacture comprising a container and
 CC the composition of matter contained within the container; (12) methods of
 CC inhibiting the growth of a cell that expresses the above protein, where
 CC the growth of the cell is at least in part dependent upon a growth
 CC potentiating effect of the above protein; (13) a method of
 CC therapeutically treating a mammal having a cancerous tumour comprising
 CC cells that express the above protein; (14) a method of determining the
 CC presence of a protein in a sample suspected of containing the protein
 CC described above; (15) methods of diagnosing the presence of a tumour in a
 CC mammal; (16) a method for treating or preventing a cell proliferative
 CC disorder associated with increased expression or activity of the above
 CC protein; and (17) a method of binding an antibody, oligopeptide or
 CC organic molecule to a cell that expresses the protein described above.
 CC The TAT sequences have cytostatic activities, and can be used in gene
 CC therapy. The composition and methods are useful for diagnosing,
 CC preventing or treating cancer. The composition is also used for preparing
 CC a medicament for the therapeutic treatment or diagnostic detection of a
 CC cell proliferative disorder or cancer. The present sequence represents a
 CC human TAT cDNA sequence from the present invention.
 XX
 XX Sequence 2186 BP; 478 A; 624 C; 673 G; 411 T; 0 U; 0 Other;
 SO
 Query Match 34.0%; Score 204.2; DB 13; Length 2186;
 Best Local Similarity 59.3%; Pred. NO. 1.3e-49;
 Matches 347; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

Db 272 CGCAAGAGGCCCAAGAAAGATGCTCAAGAACCTGTGAGGCGCTGAAGCTGGACATG 331
 Qy 257 CTTTATAGAAATATATAGTTTAAATCAAGATGACCTAGCATTTGAGAAATTTAGAG 316
 Db 332 CTGCTGCTGGGAGACAGCTGGAGCGGAGAGAGAGCTGCTGCGCGCTTCCGCGAC 391
 Qy 317 AAAGTTATCATGCTGTGTATAGACCCGTGTAGATTTTCCATAGGTGATTTACTTTGAC 376
 Db 392 CGGCGCGCTGCTGTGCGATGACCGCGCTGAGCTTCCACAGGTGAGCTTCACTTGAC 451
 Qy 377 CGGAATGTGTTATTCAGAGGCTGTTAAATGAATGACAGAGATGTCACCAATCATTCAG 436
 Db 452 CGGCGCTGCTGTGCGCGCGCGCTGCTGAGTGTCCGACCTGCTGACCGAGCGCTGGT 511
 Qy 437 CGCAGCTCACTGTCGCAAGTACATGACGCGGTTAATATGCTTTGATCATTTTTCAGAT 496
 Db 512 CCCAGCTGACCGCGCAAGTCCACGCGCGCATCAACAGCTGTTCGCGACCTAGCGGAC 571
 Qy 497 TGTGAATTTTGGCTGCTGTGTATATCTTTTGGCAATTTTAAACCCACTTACAAA 556
 Db 572 TGGAACTTCTGTGCTGTGCTCTACGCGCCGCGAGCCTTACCGCTTCCACTTGGCAGG 631
 Qy 557 CTATGATGATGATATCAAAATGTTGATGATGATGATGATGATGATGATGATGATGAT 601
 Db 632 ATCTGGAAGGCTGCGCGCGATGCTGACGAGGCGACCTTCTGA 676
 RESULT 26
 ABS78742
 ID ABS78742 standard; DNA; 619 BP.
 XX
 AC ABS78742;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE DNA encoding human NOXA17b protein.
 XX
 KW Human; NOXA; human disease; NOXA-associated disorder; cancer; addiction;
 KW Hodgkin disease; Von Hippel-Lindau syndrome; Alzheimer's disease; stroke;
 KW tuberculous sclerosis; hypercalcaemia; Parkinson's disease; depression;
 KW Huntington's disease; cerebral palsy; epilepsy; Leish-Vyhan syndrome;
 KW multiple sclerosis; ataxia-cerebellar ataxia; leukodystrophy; anxiety; pain;
 KW obesity; Crohn's disease; osteoporosis; inflammatory bowel disease;
 KW infertility; inflammatory bowel disease; atherosclerosis; hypertension;
 KW scleroderma; haemophilia; diabetes; pancreatitis; autoimmune disease;
 KW acthna; arthritis; immunodeficiency; HIV; viral infection; neurogenesis;
 KW bacterial infection; parasitic infection; graft-versus-host disease;
 KW cell differentiation; cell proliferation; haematopoiesis; wound healing;
 KW angiogenesis; gene; de.
 XX
 XX Homo sapiens.
 OS
 XX
 PN WO200272770-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 08-MAR-2002; 2002WO-US007283.
 XX
 PR 08-MAR-2001; 2001US-0274281P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275235P.
 PR 13-MAR-2001; 2001US-0275579P.
 PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 20-MAR-2001; 2001US-0277239P.
 PR 20-MAR-2001; 2001US-0277327P.
 PR 20-MAR-2001; 2001US-0277338P.
 PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.

PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-MAY-2001; 2001US-0288148P.
 PR 31-MAY-2001; 2001US-0294821P.
 PR 31-OCT-2001; 2001US-0335302P.
 PR 04-DEC-2001; 2001US-0338375P.
 PR 07-MAR-2002; 2002US-00094466.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Spytek KA, Vernet CA, Tchernev VT, Malyankar UM, Gerlach VL,
 PI Li L, Zernusen BD, Patutajan M, Gusev VY, Kekuda R, Pena CHA;
 PI Zhong M, Gangoli EA, Taupier RJ;
 XX
 DR WPI: 2002-713508/77.
 DR P-PSDB; ABG97498.
 XX
 PT New NOVX polypeptides and polymucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders, e.g. diabetes, multiple
 PT sclerosis, atherosclerosis, cancer, infections, osteoporosis or
 PT Parkinson's disease.
 XX
 PS Claim 22; Page 139; 266pp; English.
 XX
 CC The present invention relates to a new polypeptide (NOVX). The NOVX
 CC polypeptide, nucleic acid and antibody are useful in the manufacture of a
 CC medicament for treating a syndrome associated with a human disease,
 CC preferably a NOVX-associated disorder. The NOVX nucleic acids,
 CC polypeptides and antibodies are useful for treating, preventing or
 CC diagnosing diseases such as cancers, Hodgkin disease, Von Hippel-Lindau
 CC syndrome, Alzheimer's disease, stroke, tuberculous sclerosis,
 CC hypercalcaemia, Parkinson's disease, Huntington's disease, cerebral
 CC palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-
 CC telangiectasia, leukodystrophies, addiction, anxiety, depression, pain,
 CC obesity, Crohn's disease, osteoporosis, inflammatory bowel disease,
 CC infertility, inflammatory bowel disease, atherosclerosis, hypertension,
 CC scleroderma, hemophilia, diabetes, pancreatitis, autoimmune disease,
 CC asthma, arthritis, immunodeficiencies, HIV, viral, bacterial or parasitic
 CC infections, or graft-versus-host disease. The nucleic acids and small
 CC polypeptides may also be used as targets for the identification of small
 CC molecules that modulate or inhibit e.g. neurogenesis, cell
 CC differentiation, cell proliferation, haematopoiesis, wound healing and
 CC angiogenesis, in gene therapy, in generation of antibodies that bind
 CC immunospecifically to NOVX substances for use in therapeutic or
 CC diagnostic methods. The nucleic acids are further used as hybridisation
 CC probes, in chromosome mapping, tissue typing, preventive medicine, and
 CC pharmacogenomics. The present nucleic acid sequence encodes a human NOVX
 CC protein of the invention
 CC
 XX
 SQ Sequence 619 BP; 97 A; 219 C; 207 G; 96 T; 0 U; 0 Other;
 XX
 Query Match 33.8%; Score 203.4; DB 6; Length 619;
 Best Local Similarity 60.0%; Pred. No. 1.2e-49;
 Matches 339; Conservative 0; Mismatches 226; Indels 0; Gaps 0;
 XX
 QY 37 GGGCAGATGTTCTTAATTCAGAAAACCTGGCGTTCAGGCAAGAAAGTCTTGGG 96
 DB 4 GGGCAGATGACCTTCAGACCAAGAGCTTGCTGCGAGGCCAGAGAGAGCTCTGAG 63
 QY 97 TAAATGTGTCTCCAAATTCATCGCACCACTTATATAGACAGACAAAGTAGAGGTGCT 156
 DB 64 TAAAGTGGCGTCCAGAGCAGTGTGGCCGTGCTGTGAGATGACACAGCAGTAGGTGCT 123
 QY 157 GGATGAGCTCTACAGAGTACCCAGGAGTACCCCAAAACAGAGAGGCGAGAGAGT 216
 DB 124 GGAATGAGCTGTACCGGCAACAGGAGTTCAGCCAGCCGCAAGAGGCGCCAGAGAT 183
 QY 217 CATCAAGAGCCTCATCAAGAGCAGTCAATCAAGCTGGCATTCCTTATAGAGAAATAGT 276
 DB 184 GCTCAAGAACTGTGTAGAGGTGCGCTGAGGCTGGAGCTGCTGCTGGGGACCAAGCT 243
 QY 277 TATCAAGATGAGCTGATGATGAGAGAAATTAAGAGAAAGTTATCATCAGCTTCTAT 316
 || ||||| || ||||| || ||||| || ||||| || ||||| || ||||| || |||||

DB 244 GGGCGATGAGAGAGTGGCGCTGCTGCGGCGCTTCCGCCACGCGGCGCGCTGAGCCAT 303
 QY 337 GACCGTGTCAGATTTCATCAGTGTGATTAATCCTTTGACCGGAAATGCTTATCCAGGT 396
 DB 304 GACGGCGCTGACCTTCACAGGTGATCTTCACTTCAGACCGCGCGTGGCGCGCG 363
 QY 397 GTTAATGATGATGAGAGATGCTGACCAATCATTCATAGCGCCACCTCATGCGCAAGTC 456
 DB 364 GCTGCTCAGATGCTCCGACACTCTGTCCACAGGCGCTGGGTCCCACTGACCGCCAGTC 423
 QY 457 ACATGAGCGGATTAATTAATGCTTGTATCATTTTCAATTTGATGAAATTTGGTGCCTT 516
 DB 424 CCACGCGCGCATCAACACAGTGTGGGCCACCTTACCGACATGCGACTTCTGTCGCGCT 483
 QY 517 GTTATATCTTTTGGGAATTTTAAACCCCATTTAAMAAACATATGATGTATCAACA 576
 DB 484 CTACGCGCGCGCGCGCGCGCTTACCGCTCCACCTGCGAGATTCGAGAGGCGTGGCGG 543
 QY 577 AATGTTGATGAGAGAGACATATGA 601
 DB 544 GATGCTGAGACGAGCGAGCTCTGA 568
 XX
 RESULT 27
 ABK34943
 ID ABK34943 standard; cDNA; 1113 BP.
 XX
 AC ABK34943;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human cDNA encoding secreted protein #81.
 XX
 KW Human; secreted protein; gene; ss; nutritional supplement; haemophilia;
 KW viral infection; bacterial infection; fungal infection; diabetes; asthma;
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
 KW tissue regeneration; wound healing; burn; haematopoiesis;
 KW myeloid cell deficiency; lymphoid cell deficiency.
 XX
 OS Homo sapiens.
 XX
 PN MO200177288-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US010224.
 XX
 PR 06-APR-2000; 2000US-0195582P.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Wong GG, Clark HP, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulakota K, Graham JR;
 XX
 DR WPI: 2002-179321/23.
 XX
 PT Five hundred and ninety two polymucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for treating
 PT immune deficiencies and disorders such as autoimmune disorders.
 XX
 PS Claim 1; Page 109; 372pp; English.
 XX
 CC The invention relates to 592 polymucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins. The polymucleotides can be used as probes for the
 CC identification and isolation of full length cDNA and genomic DNA. The
 CC polymucleotides and proteins can also be used as nutritional supplements.
 CC The proteins are useful in the treatment of various immune deficiencies
 CC and disorders such as viral infections, bacterial infections, fungal
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple

CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
 CC and conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment of
 CC burns, incisions and ulcers. The proteins are also useful for regulating
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
 CC Sequences ABK34863-ABK35454 represent polynucleotides of the invention
 XX
 XX Sequence 1113 BP; 282 A; 311 C; 304 G; 216 T; 0 U; 0 Other;

Query Match 30.3%; Score 182; DB 6; Length 1113;
 Best Local Similarity 58.7%; Pred. No. 3.6e-43;

Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGATGCGCAGATGCTTTAATCCAAAACCTGGCGCTTCAGGCAAAAGAGA 89
 DB 52 CTCAGAGACCATGAGATCCTTCAGCTCAAGAGCCTGGCACTGCAAGAGAGAGAC 111
 QY 90 TCTTGGGTAAATGCTGTCCAAATCCATGCGCCACCACTTAATGACGACAGATAGTG 149
 DB 112 TACTGAGTAAAGATGGCGGTGCTGTGCTCATCTTCATAGATGAGACAAGCAGTG 171
 QY 150 AGGTGCTGATGAGCTCTACAGAGTACCGAGGAGTACACCAAAACAGAGAGGACG 209
 DB 172 AGGTGCTGATGAGCTCTACAGAGTACCGGTGCTCAAGAGTACAGCAGCGGCGCCAGGCC 221
 QY 210 AGAAGATCATCAAGAACTCATCAAGACAGTCAATCAAGCTGCGCATTTCTTATAGAAATA 269
 DB 232 AGCGGCTGATCAAGAGACCTGATCAAGAGTGCATCAAGTGGCTGTGCTGACCGCAATG 291
 QY 270 ATCAGTTTAAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTTCATCAGC 329
 DB 292 GCTCTCTGGGCCCCAGTGAAGCTGCGCTGAGCTACCGGCTTTCGCCAAGAGCTCGGAGG 351
 QY 330 TTGCTATGACCGGTGCTGATTTCCATCAGGTGATTAACCTTTGACCGGAATGTGTAT 389
 DB 352 GTGCCATGAGCGGACCTTAACTTTGGTGAAGTAACTTCACTTGAAGGCTGTCTGTG 411
 QY 390 CCAGGCTGTTAAATGATGAGAGATGCTGACCAACCAATTCATTCAGCGGCACTTCAGT 449
 DB 412 CTGGCTGTGACCGGATGCGGAGATGTCTCTAAGAGTTGGTGAACACCACTTCAAGC 471
 QY 450 CCAAGTCAATGAGACGGGTTAATATGCTTTGATCATATTTTTCAGATTGTGAATTTTGG 509
 DB 472 CCAAGTCAATGAGACGGGTTAATATGCTTTGATCATATTTTTCAGATTGTGAATTTTGG 511
 QY 510 CTGGCTGTGATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATGTGATGTA 569
 DB 532 CGGCGCTTATGAGGCT-----GACTTCACTCAGACCACTTGGCAAGATCTGTGAGCGAC 585
 QY 570 TCACAAATATGTTGATGAGAGACATATGA 601
 DB 586 TCAGAAAGCTCTAAGAGAGAGAGAGCTTGA 617

RESULT 28
 AAD38699
 ID AAD38699 standard; cDNA; 1165 BP.
 XX
 AC AAD38699;
 XX
 DT 23-SEP-2002 (first entry)
 XX
 DB Human LP237 secreted protein encoding cDNA.
 XX
 KM Human; secreted protein; atherosclerosis; Alzheimer's disease; LP237;
 KM diabetic retinopathy; severe combined immunodeficiency; pancreatitis;
 KM rheumatoid arthritis; colorectal adenoma; haemolytic anaemia; cancer;
 KM reperfusion injury; arteriosclerosis; wound healing; transgenic animal;
 KM gene therapy; neoplasm; transgenic; psoriasis; ischemia; carcinoma;
 KM gene, ss.

XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FH CDS 100..654
 FT /tag= a
 FT /product= "human LP237 secreted protein"
 FT sig_peptide 100..180
 FT /tag= b
 FT mat_peptide 181..651
 FT /tag= c
 FT /product= "Mature human LP237 secreted protein"
 XX
 XX WO200226801-A2.
 XX
 XX 04-APR-2002.
 XX
 XX 14-SEP-2001; 2001WO-US026026.
 XX
 XX 28-SEP-2000; 2000US-0236088P.
 XX
 XX (ELIL) LILLY & CO ELI.
 XX
 XX Su EW, Wang H;
 XX
 XX WPI; 2002-471259/50.
 XX
 XX P-PSDB; AAE23984.
 XX
 XX Claim 1; Page 140-141; 145pp; English.

The invention relates to human secreted polypeptides designated LP095, LP191, LP217, LP220, LP221, LP222, LP229, LP237 or LP238 and nucleic acid molecules encoding such polypeptides. Novel secreted proteins of the invention are used for treating diseases such as atherosclerosis, Alzheimer's disease, diabetic retinopathy, psoriasis, pancreatitis, arteriosclerosis, rheumatoid arthritis, colorectal adenomas, severe combined immunodeficiency, ischemia, carcinoma, haemolytic anaemia, reperfusion injury, neoplasms and cancer especially liver cancer. They are also used for wound healing. Polynucleotides of the invention can be used to generate transgenic animals or knock out animals, which in turn, are useful in the development and screening of therapeutically useful CC reagents for use in the treatment of diseases associated with LP CC polypeptide associated activity. They are also used in gene therapy. The CC present sequence is human LP237 secreted protein encoding cDNA

Sequence 1165 BP; 305 A; 319 C; 312 G; 229 T; 0 U; 0 Other;

Query Match 30.3%; Score 182; DB 6; Length 1165;
 Best Local Similarity 58.7%; Pred. No. 3.7e-43;

Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGATGCGCAGATGCTTTAATCCAAAACCTGGCGCTTCAGGCAAAAGAGA 89
 DB 89 CTCAGAGACCATGAGATCCTTCAGCTCAAGAGCCTGGCACTGCAAGAGAGAGAC 148
 QY 90 TCTTGGGTAAATGCTGTCCAAATCCATGCGCCACCACTTAATGACGACAGATAGTG 149
 DB 149 TACTGAGTAAAGATGGCGGTGCTGTGCTCATCTTCATAGATGAGACAAGCAGTG 208
 QY 150 AGGTGCTGATGAGCTCTACAGAGTACCGAGGAGTACACCAAAACAGAGAGGACG 209
 DB 209 AGGTGCTGATGAGCTCTACAGAGTACCGAGGAGTACACCAACCGGCGCCAGGCC 268
 QY 210 AGAAGATCATCAAGAACTCATCAAGACAGTCAATCAAGCTGCGCATTTCTTATAGAAATA 269
 DB 269 AGCGGCTGATCAAGAGACCTGATCAAGAGTGCATCAAGTGGCTGTGCTGACCGCAATG 328
 QY 270 ATCAGTTTAAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTTCATCAGC 329

Db 329 GCTCCTTTGGCCCACTGAGCTGGCCCTGGCTACCGCTTTGCGCAGAAAGTCCGGCAGG 388
 QY 330 TTGCTATGACCGGTGTCAGTTTCCATCAGGTGATATATACCTTTGACCGGAATGTGTAT 389
 Db 389 GTGCCATGACGGACTTAACTTTGGAGGTAACTTACCTTGAGGCTGTCTGTCTGG 448
 QY 390 CCAGGCTGTTAAATGATGACAGAGATGCTGACCAAAATCATTCACGCCCACTCACTG 449
 Db 449 CTGGCTGTGACCGGATGCGGGGATGTCTGTAAAGTTGGTGAACACACCTCAACG 508
 QY 450 CCAAGTCACATGACGCGGTTAATATGCTTTGATCATATTTTCAATTTGAAATTTTGG 509
 Db 509 CCAAGTCACATGACGCGGATGCGGATGCTGTGTATCATCTCTGACCCAGGCTCTCTCA 568
 QY 510 CTGCTTGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGATGATGA 569
 Db 569 CGGCTCTTATGAGGCTCT-----GACCTTCACTCAGACCTTGGCAAGATCTGTGAGCAG 622
 QY 570 TCACAAATGTTGATGAGAGACATATGA 601
 Db 623 TCAGGAAGCTGCTAAGACGAAGGAAGCTCTGA 654

RESULT 29
 AB211655
 ID AB211655 standard; cDNA; 1175 BP.
 AC AB211655;
 DT 20-JUN-2003 (first entry)
 DE Human polynucleotide SEQ ID NO 537.
 XX
 KW Human; genome mapping; gene therapy; food supplement; virus; fungus;
 cell-proliferative disorder; neurodegenerative disease; bacterial;
 Parkinson's disease; Alzheimer's disease; autoimmune disease;
 multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 arthritis; cytotoxic; immunomodulator; neotropic; neuroprotective;
 anti-inflammation; antidiabetic; immunosuppressive; dermatological;
 haemostatic; vulnerary; fungicide; antibacterial; viral; protozoicide;
 antitubercular; gene; ss.
 KW
 OS Homo sapiens.
 PN MO200270539-A2.
 XX
 PD 12-SEP-2002.
 XX
 PF 05-MAR-2002; 2002MO-US005095.
 XX
 PR 05-MAR-2001; 2001US-00799451.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue Y, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX
 DR WPI; 2002-759812/82.
 DR P-PSDB; ABP69438.
 XX
 PT New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or platelet
 PT or coagulation disorders.
 XX
 PS Claim 1; SEQ ID NO 537; 1012bp + Sequence Listing; English.
 CC The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences (AB211119-
 CC AB212066) or their mature protein coding portion, active domain coding
 CC protein or complementary sequences. The polynucleotides are useful for
 CC identifying expressed genes or for physical mapping of human genome. The

CC encoded polypeptides (ABP68902-ABP69849) are useful as molecular weight
 CC markers, as a food supplement, for generating antibodies, in medical
 CC imaging, screening and diagnostic assays and for treating cell-
 CC proliferative disorders (cancer), neurodegenerative diseases (Parkinson's
 CC or Alzheimer's disease), autoimmune diseases (multiple sclerosis,
 CC diabetes, lupus) genetic disorders, myeloid or lymphoid disorders,
 CC platelet or coagulation disorders, wound, burns, incision, ulcers, liver
 CC or lung fibrosis, infections (bacterial, viral, fungal, parasitic),
 CC arthritis, etc. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIP0 at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 1175 BP; 306 A; 324 C; 313 G; 232 T; 0 U; 0 Other;

Query Match 30.3%; Score 182; DB 6; Length 1175;
 Best Local Similarity 58.7%; Pred. No. 3.7e-43;
 Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGATGGCCACAGATGCTTTAATTCAAAACCTGGCCGTTACAGGACAAAAGAGA 89
 Db 102 CTCGAGACCCATGAGATCCTTCAAGCTCAAAAGCTGGCACTGCAAGAGAAAGAGC 161
 QY 90 TCTTGGTAAATGCTGTCCAATCCATCGCCACCACTTAATGACGACAGATAGTG 149
 Db 162 TACTGATGTAAGTGGCGGGTCCGCTCTGTGCTCATCTCTCATATGATGAGACAGACAGTG 221
 QY 150 AGGTGCTGATGATGCTCTACAGAGTACAGGAGGATACCCAAAACAAAGAGGCGAG 209
 Db 222 AGGTGCTGATGATGCTCTACAGGATGATCCAAAGAGTACAGGACGCGGCCCGCCAGGCC 281
 QY 210 AGAAGTCAATCAAGAACTCATCAAGACAGTCAAGCTGCGCATTTCTTTATAGAAATA 269
 Db 282 AGCGGTATATCAAGACCTGATCAAGTCAAGTCCATCAAGTGGCTGTGTCGACCGGAAG 341
 QY 270 ATCAGTTAATCAAGATGAGCTAGCATGATGATGAGAAATTAAAGAAAGTTCAATCAGC 329
 Db 342 GCTCCTTTGGCCCCAGTGAAGTGGCCCTGCTACCCGCTTTGCCAGAAAGCTGCCGACAG 401
 QY 330 TTGCTATGACCGGTGTCAGTTTCCATCAGGTGATATATACCTTTGACCGGAATGTGTAT 389
 Db 402 GTGCCATGACGGACTTAACTTTGGAGGTAACTTACCTTCAAGGCTGTCTGTCTGG 461
 QY 390 CCAGGCTGTTAATGATGACAGAGATGCTGACCAAAATCATTCAGGCCCACTCACTG 449
 Db 462 CTGGCTGTGACCGGATGCGGGGATGTCTGTAAAGTTGGTGAACACACCTCAACG 521
 QY 450 CCAAGTCACATGACGCGGTTAATATGCTTTGATCATATTTTCAATTTGAAATTTTGG 509
 Db 522 CCAAGTCACATGACGCGGATGCGGATGCTGTGTATCATCTCTGACCCAGGCTCTGCA 581
 QY 510 CTGCTTGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGATGATGA 569
 Db 582 CGGCTCTTATGAGGCTCT-----GACCTTCACTCAGACCTTGGCAAGATCTGTGAGCAG 635
 QY 570 TCACAAATGTTGATGAGAGACATATGA 601
 Db 636 TCAGGAAGCTGCTAAGACGAAGGAAGCTCTGA 667

RESULT 30
 ADMA4173
 ID ADMA4173 standard; cDNA; 1175 BP.
 AC ADMA4173;
 DT 03-JUN-2004 (first entry)
 DE Novel human arginine-rich protein cDNA #537.
 XX
 KW ss; gene; human; arginine-rich protein; cancer; inflammation;
 KW genetic disorder.
 XX
 OS Homo sapiens.

```
XX US2004053250-A1.
BN
PD 18-MAR-2004.
PF 21-NOV-2002; 2002US-00302172.
PR 05-MAR-2001; 2001US-00799451.
PR 05-MAR-2002; 2002WO-US005095.
PR 20-AUG-2002; 2002US-00225251.
XX
XX (TANG/) TANG Y T.
XX (XUEA/) XUE A.
XX (DRMA/) DRMANAC R T.
XX
XX Tang YT, Xue A, Drmanac RT,
XX
XX WPI; 2004-238579/22.
XX
PT New isolated arginine-rich protein-like polynucleotides and polypeptides,
PT useful for diagnosing and/or treating conditions associated with aberrant
PT activity of the arginine-rich polypeptides, such as cancer and
PT inflammation.
XX
XX Disclosure; SEQ ID NO 537; 51pp; English.
XX
XX The invention relates to an isolated polynucleotide. The methods and
XX compositions of the present invention are useful for the diagnosis and/or
XX treatment of diseases or conditions associated with aberrant expression
XX or activity of the arginine-rich protein-like polypeptides, such as
XX cancer and inflammation. They can also be used in forensics, gene
XX mapping, identification of mutations responsible for genetic disorders,
XX and in assessing biodiversity. The present sequence represents a novel
XX human arginine-rich protein cDNA.
XX
XX Sequence 1175 BP; 306 A; 324 C; 313 G; 232 T; 0 U; 0 Other;
XX
XX Query Match 30.3%; Score 182; DB 12; Length 1175;
XX Best Local Similarity 58.7%; Pred. No. 3.7e-43;
XX Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;
XX
QY 30 CTCGAGTGGCAGACAGATGCTTTAATTCAGAAACCTGGCCCTTCAGGCACAAAGAAG 89
DB 102 CTCAGAGACCCATGAGTCTTCTAGCTCAAGAGCCCTGCACCTGCAGACAGAGAAGAGC 161
QY 90 TCTTGGGTAAATGATGCTCAATTCATCGCCACCACTTAATAGACAGACAAGTAGT 149
DB 162 TACTGATGAAGATGGGGGCGCTGTGTGCTCATCTCTTCATAGATGAGACAAGCAGT 221
QY 150 AGGTGCTGATGAGTCTTACAGAGTACCAAGGAGTACACCAAAACAAAGAGAGCGAG 209
DB 222 AGGTGCTGATGAGTCTTACAGAGTACCAAGGAGTACACCAAGGAGAGCGAGCGCC 281
QY 210 AGAAGTATCAAGAACCTCATCAAGACATCAAGCTGGCCATCTTTATAGAAAT 269
DB 282 AGGCGGTGATCAAGACCTGATCAAGTGCATCAAGGCGTGTGCTGCACCGGAAGT 341
QY 270 ATCAGTTTATCAAGATGAGTCAAGTTCAGTGAAGAAATTAAGAAGATTCATCAGC 329
DB 342 GCTCCTTTGGCCCCCATGAGTGGCCCTGCTACCCGCTTCCCGAAGAGTCCGGCAG 401
QY 330 TTGCTATGACCGTGTGCTTCCATCAGTGTGATTAATCTTTGACCGGATGTGTAT 389
DB 402 GTGCCATGACGCGCATTAAGCTTTGTGAGGTATCACTTCGAGGCTGTGTTCTGG 461
QY 390 CAGGCGTGAATTAATGATGAGAGATGCTGACCAAAATCATTCAGCGCCACTCAGT 449
DB 462 CTGGCTGTGACCGAGTGCAGGAGTGTGCTAGAGTTGTGGAACACACCTCAGC 521
QY 450 CCAAGTCACATGACGCGGTAAATATGCTCTTTGATCATTTTTCAGATTGTGAATTTTGG 509
DB 522 CCAAGTCACATGACGCGCGCATCCGCAAGTGTGTTGATCATCTTCTGACCCAGGCTCTCA 581
```

```
QY 510 CTGCTGTGATATATCTTTTGGGAATTTAAACCCGACTTACAAAGACTATGATGATGTA 569
DB 582 CGGCTCTTATGGGCTT-----GACTTCACTAGACCTTGGCAAGATCTGTAGCGAC 635
QY 570 TCACAAAAATGTTGATGAGAGACATATGA 601
DB 636 TCAGAAAGCTGCTAGACGAAGGAGAGACTCTGA 667

RESULT 31
AAF59594
ID AAF59594 standard; cDNA; 1268 BP.
XX
XX AAF59594;
AC
XX
XX 24-APR-2001 (first entry)
DT
XX
XX Human cell cycle and proliferation protein CCYPR-5 cDNA, SEQ ID NO:59.
DE
XX
XX Cell cycle and proliferation protein; CCYPR; human; agonist; antagonist;
XX gene therapy; detection; gene therapy; transgenic animal disease model;
XX immune disorder; developmental disorder; cell signalling disorder;
XX cell proliferative disorder; cancer; tumour; anaemia; epilepsy;
XX arteriosclerosis; asthma; allergy; diabetes mellitus;
XX menstrual cycle disorder; bacterial infection; ss.
XX
XX Homo sapiens.
XX
XX MO200107471-A2.
XX
XX 01-FEB-2001.
XX
XX 21-JUL-2000; 2000WO-US019948.
XX
XX 21-JUL-1999; 99US-0145075P.
XX 08-SEP-1999; 99US-0153129P.
XX 10-NOV-1999; 99US-0164647P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
XX Hillman JL, Lal P, Tang YT, Yue H, Au-Young J, Bandman O;
XX Azimzai Y, Yang J, Lu DM, Baughn MR, Patterson C, Shah P;
XX P-PSDB; AAB60457.
XX
XX WPI; 2001-112727/12.
XX
XX Human cell cycle and proliferation proteins and polynucleotides are used
XX to treat, diagnose and prevent immune, developmental and cell signalling
XX disorders and cell proliferative disorders including cancer.
XX
XX Claim 5; Page 171; 205pp; English.
XX
XX Sequences AAF59590-AAF59643 represent cDNAs encoding 54 human cell cycle
XX and proliferation proteins (CCYPR), AAB60453-AAB60506. CCYPR and agonists
XX of CCYPR are used to treat diseases or conditions associated with
XX decreased expression of functional CCYPR, while CCYPR antagonists are
XX used to treat diseases or conditions associated with overexpression of
XX functional CCYPR. Monoclonal or polyclonal antibodies to CCYPR may be
XX used in enzyme-linked immunosorbent assays (ELISA) or radioimmunoassays to
XX detect CCYPR. CCYPR itself may be used to detect compounds e.g.,
XX antibodies, oligonucleotides and proteins (receptors) that specifically
XX bind to CCYPR, and in drug screening methods to identify compounds that
XX modulate the activity of CCYPR. CCYPR nucleotides can be used to generate
XX transgenic animal models of human disease, and can be used in gene
XX therapy in target cells with genetic abnormalities with respect to the
XX expression of CCYPR for the treatment or prevention of a disorder
XX associated with CCYPR. Diseases which can be diagnosed, treated and
XX prevented using CCYPR proteins, nucleic acids, agonists or antagonists
XX include immune, developmental and cell signalling disorders, and cell
XX proliferative disorders including cancer. Specific examples of these
XX disorders include anaemia, epilepsy, arteriosclerosis, asthma, cancer,
XX allergies, diabetes mellitus, disorders of the menstrual cycle and
XX infections caused by bacteria
```


XX Sequence 1268 BP; 343 A; 345 C; 350 G; 230 T; 0 U; 0 Other;
 SQ Query Match 30.3%; Score 182; DB 4; Length 1268;
 Best Local Similarity 58.7%; Pred. No. 3.9e-43;
 Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;
 QY 30 CTCCTGAGGACCAAGATGCTTAAATCCAAAACCTGGCCGCTGAGGACAAAAGAAAGA 89
 DB 92 CTCAGAGACCCATGAGTCTTCACTCAAGAGCTGGCACTGCAAGACAGAAAGAAC 151
 QY 90 TCTTGGGTAATATGCTGCAATTCATGCGCCACCACTTAATAGACACAAAGTAGT 149
 DB 152 TACTAGTAGAATGAGGGGTGCTGCTGTGCTATCTTCAATAGTAGACAGAGAGT 211
 QY 150 AGGTGCTGATGAGCTCTACAGAGTACCAAGGAGTACACCCAAAACAGAGAGGACG 209
 DB 212 AGGTGCTGATGAGCTCTACAGAGTACCAAGGAGTACACGACAGCCGCGCCAGGCCC 271
 QY 210 AGAAGATCATCAAGACCTCATCAAGACGTCAAGCTGCGCATTTTAAAGATA 269
 DB 272 AGGCGCTGATCAAGAGCTGATCAAGAGTCAAGAGTGTGCTGACCGCAATG 331
 QY 270 ATCAGTTTAATCAAGATGAGTACATGATGAGAAATTTAAGAAAGATTATCAGC 329
 DB 332 GCTCCTTGGCCCAATGAGCTGCGCTGCTACCCGCTTGGCCAGAGCTGCGGACG 391
 QY 330 TTGCTATGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 389
 DB 392 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
 QY 390 CCAAGCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 449
 DB 452 CTGGCTGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 511
 QY 450 CCAAGCTGATGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATG 509
 DB 512 CCAAGCTGATGAGCGGCTGATGATGATGATGATGATGATGATGATGATGATG 571
 QY 510 CTGGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 569
 DB 572 CCGCTGCTGATGAGCGCTGATGATGATGATGATGATGATGATGATGATGATG 625
 QY 570 TCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
 DB 626 TCAGGAGCTGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 657
 RESULT 32
 ACH76016
 ID ACH76016 standard; DNA; 500 BP.
 XX ACH76016;
 XX 29-JUL-2004 (first entry)
 XX Human genome derived single exon probe #9211.
 XX Human; probe; ss; gene expression; single exon probe; microarray;
 XX alternative splicing event; genomic alteration.
 XX Homo sapiens.
 OS US2003194704-A1.
 PN 16-OCT-2003.
 PD 03-APR-2002; 2002US-00029386.
 PF 03-APR-2002; 2002US-00029386.
 PR 03-APR-2002; 2002US-00029386.
 XX (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 XX

PA (HANZ/) HANZEL D K.
 XX Penn SG, Rank DR, Hanzel DK;
 PT WPI; 2004-119264/12.
 PT New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX Claim 15; SEQ ID NO 9211; 80bp; English.
 PS The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC a method of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterizing
 CC alternative splicing events, in detecting and characterizing gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?DocID=20030194704
 XX
 SQ Sequence 500 BP; 164 A; 71 C; 107 G; 158 T; 0 U; 0 Other;
 Query Match 28.0%; Score 168.4; DB 12; Length 500;
 Best Local Similarity 99.4%; Pred. No. 2.6e-39;
 Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 432 TTCAGGCGACCTCACTGCGCAGTGCATGAGCGGCTTAATATGCTTGAATCTTT 491
 DB 1 TTCAGGCGACCTCACTGCGCAGTGCATGAGCGGCTTAATATGCTTGAATCTTT 60
 QY 492 CAGATTGTAATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 551
 DB 61 CAGATTGTAATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
 QY 552 AAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
 DB 121 AAAAATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 170
 RESULT 33
 AAC10352
 ID AAC10352 standard; cDNA; 340 BP.
 XX AAC10352;
 AC

XX 06-OCT-2000 (first entry)
DT Human secreted protein 5' EST, SEQ ID NO: 14427.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX Gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX BP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-00200610.
XX 26-FEB-1999; 99US-0122487P.
XX (GSEST) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 14427; 71bp + Sequence Listing; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNA encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or poly(A) RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 340 BP; 92 A; 79 C; 85 G; 79 T; 0 U; 5 Other;
Query Match 26.5%; Score 159.2; DB 3; Length 340;
Best Local Similarity 98.8%; Pred. No. 1.2e-36;
Matches 158; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 45 ATGCTCTTAATTCACAAAACCTGGCGGTTCAGGCACACAAAAGATCTTGGAATAATG 104
DB 181 ATGCTCTTAATTCACAAAACCTGGCGGTTCAGGCACACAAAAGATCTTGGAATAATG 240
QY 105 TGTCCAAATTCATCGCCACCACTTAATAGACACACAACTAGTAGGTCTGATGAC 164
DB 241 TGTCCAAATTCATCGCCACCACTTAATAGACACACAACTAGTAGGTCTGATGAC 300
QY 165 TCTACAGATGACACAGGAGTACACCCAAACAAAGAGA 204
DB 301 TCTACAGATGACACAGGAGTACACCCAAACAAAGAGA 340
RESULT 34
AB107183
ID AB107183 standard; cDNA; 829 BP.
XX
AC AB107183;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 16031.

XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
XX 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PDB; ABB63080.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO: 16031; 21bp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins (AAB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 829 BP; 229 A; 208 C; 210 G; 182 T; 0 U; 0 Other;
Query Match 24.3%; Score 146.2; DB 4; Length 829;
Best Local Similarity 57.6%; Pred. No. 1.3e-32;
Matches 262; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
QY 35 ATGGCCACAGATGTCTTAAATTCACAAAACCTGGCGGTTCAGGCACACAAAAGATCTTG 94
DB 263 ATGGCCACAGATGTCTTAAATTCACAAAACCTGGCGGTTCAGGCACACAAAAGATCTTG 322
QY 95 GGTAAATAGTGTCCAAATTCATCGCCACCACTTAATAGACACACAACTAGTAGGTG 154
DB 323 TCACGATGCGCACAAAACATCGCCACCACTTAATAGTAGGACGACGCGGTGCGTG 382
QY 155 CTGATGAGCTCTACAGAGTACACAGGAGTACACCCAAACAAAGAGAGGACGAGAG 214
DB 383 CTGACCAACCTCTACAGGCTCTGCAAGATGACACAGGAGACAAAGCGGAGGAGAG 442
QY 215 ATCATCAAGAACCTATCAAGACAGTCATCAAGCTGCGCATCTTATAGGAATATCG 274
DB 443 CTGATCAAGAACCTATCAAGATGATCAAGATGCTGCTCCACCGAACAATTCAG 502
QY 275 TTTAATCAAGATGAGTACATTTAGTGAAGAAATTTAAGAAAGTTATCATGCTTGC 334
DB 503 TTCAGCGACAGAGAGTGCAGAAAGGCGAGCTCTTTAAGAGAAAGTTCAAAACAGCGAG 562
QY 335 ATGACCGTGTGATGATTTTCATCAGGTGATTTATCTTGAACCGGAATGTATTACAG 394
DB 563 CTGTCCATCATATCATTTTACAGAGTGCATTCAGGTTGACCTGCGGACGCAAAAG 622
QY 395 CTGTTAAATGATGACAGAGATGCTGCACCAATCATTCAGGCGCACTGCGCAAG 454
DB 623 TCATATGCGCAATTCGCAAGTGGCGCTGAAGTGAATTGTGACCGCATCTCACAAAG 682

Query Match 21.0%; Score 126.4; DB 3; Length 252;
Best Local Similarity 98.6%; Pred. No. 5.4e-27;
Matches 138; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTTACGGTCCGCGCGCGTGC-CGACTCTCGATGGCCAGATGCTTTAATTCCA 59
DB 113 CTTACGGTCCGCGCGCGTGC-CGACTCTCGATGGCCAGATGCTTTAATTCCA 112

QY 60 AAAACCTGGCGCTTCAGGCAAAAAGATCTTGGGTAAATGGTGTCCAAATCCATCG 119
DB 173 AAAACCTGGCGCTTCAGGCAAAAAGATCTTGGGTAAATGGTGTCCAAATCCATCG 232

QY 120 CCACCACTTAATAGACGAC 139
DB 233 CCACCACTTAATAGACGAC 252

RESULT 37
AAK56622
ID AAK56622 standard; cDNA; 441 BP.
XX
AC AAK56622;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:1682.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
PF 17-JUN-2001; 2001WO-US001354.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0220863P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234233P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234984P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241877P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.

[illegible]

CC	AAK54959-60	AAK64702	encode the human immune/haematopoietic antigen (I)	
CC	amino acid sequences given in AAM62170 to AAM91521.	(I)	have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production.	(I)		
CC	proteins and polynucleotides may be used in the prevention, diagnosis and			
CC	treatment of diseases associated with inappropriate (I) expression. For			
CC	example, they may be used to treat disorders associated with decreased			
CC	expression by rectifying mutations or deletions in a patient's genome			
CC	that affect the activity of (I) by expressing inactive proteins or to			
CC	supplement the patient's own production of (I). Additionally, (I)			
CC	polynucleotides may be used to produce the secreted (I), by inserting the			
CC	nucleic acids into a host cell and culturing the cell to express the			
CC	protein. (I) proteins and polynucleotides may be used to prevent,			
CC	diagnose and treat immune/haematopoietic-related diseases, especially			
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703			
CC	and AAK7694 represent human immune/haematopoietic antigen genomic			
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169			
CC	represent sequences used in the exemplification of the present invention			
xx				
xx	Sequence 441 BP; 99 A; 119 C; 127 G; 89 T; 0 U; 7 Other;			
Query Match	19.7%;	Score 118.4;	DB 4;	Length 441;
Best Local Similarity	59.3%;	Pred. No. 1.7e-24;		
Matches 194;	Conservative 3;	Mismatches 130;	Indels 0;	Gaps 0
Qy	39	CCACGATGTCCTTTATATCCAAAAAACCTGACCGTTGACGACACAAAAGAGATCTTGCGTA	98	
Db	78	CCATGAGATGCTTACAGCTCAAAAGAGCTTGGCACTGCAAGCAGAGAAAGAGCTACTAGTA	137	
Qy	99	AAATGCTGTCCAAATTCATGCGCACCAACCTTATATAGACACCAACAGTAGTAGGTCCTGG	158	
Db	138	AGATGGCGGATGCGCTCTGTGGCTGCTATCTCTTATATATAGAGCAACACAGTAGAGTCTGG	197	

Oy	155	ATAGAGCTCAACAGAGGACCCAGGGAGTACACCCAAAACAGAAAGAGGACAGAAAGATCA	218
Db	198	ATAGAGCTCAACCGTGTGTCCAAAGAGTACAGCACAGCCGGCCCCAGGCCCCAGGCGGTGA	257
Oy	219	TCGAAGACCTCATCAAGACAGTCATCAAGAGCTGGCCATTTCTTTATAGGAATATACGTTTA	278
Db	258	TCGAAGACCTGATCAAAAGTGGSCATCAAGAGTGGCTGTGTCTGTSCAACCGAAATGGCTCTTTG	317
Oy	279	ATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTTCATCAGCTTGCTATGGA	338
Db	318	GCCCCAGTGAAGCTGGGCCCTGGCTAACCGCTTTTCGACGAATCTCGGCAAGATGTCATGA	377
Oy	339	CCGTGTGACGTTTCCATCAGGTGATTT	365
Db	378	GGGCATTTGAGCTGTGTGAATGATTT	404

XX	RESULT 38	
XX	ABL07182/c	
XX	ID ABL07182 standard; cDNA; 4406 BP.	
XX	AC ABL07182;	
XX	DT 26-MAR-2002 (first entry)	
DE	Drosophila melanogaster expressed polynucleotide SEQ ID NO 16028.	
XX	Drosophila; developmental biology; cell signalling; insecticide;	
KW	pharmaceutical; gene; ss.	
XX	Drosophila melanogaster.	
OS	WO200171042-A2.	
PN		
PD	27-SEP-2001.	
XX		
PF	23-MAR-2001; 2001WO-US009231.	
XX		
PR	23-MAR-2000; 2000US-0191637P.	
XX	11-JUL-2000; 2000US-00614150.	
PA	(PEKE) PE CORP NY.	
PI	Venter JC, Adams M, Li PWD, Myers EW;	
DR	WPI; 2001-656660/75.	
XX	P-PsDB; ABB63079.	
PT	New isolated nucleic acid detection reagent for detecting 1000 or more	
XX	genes from Drosophila and for elucidating cell signaling and cell-cell	
XX	interactions.	
PS	Claim 1; SEQ ID NO 16028; 21pp + Sequence Listing; English.	
XX		
CC	The invention relates to an isolated nucleic acid detection reagent	
CC	capable of detecting 1000 or more genes from Drosophila. The invention is	
CC	useful in developmental biology and in elucidating cell signalling and	
CC	cell-cell interactions in higher eukaryotes for the development of	
CC	insecticides, therapeutics and pharmaceutical drugs. The invention	
CC	discloses genomic DNA sequences (AB116176-AB130511), expressed DNA	
CC	sequences (ABL01840-ABL16175) and the encoded proteins (AAB57237-	
CC	ABB72072). The sequence data for this patent did not form part of the	
CC	printed specification, but was obtained in electronic format directly	
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences	
XX		
XX	Sequence 4406 BP; 1178 A; 1021 C; 1027 G; 1180 T; 0 U; 0 Other;	
Query_Match	19.3%; Score 116; DB 4; Length 4406;	
Best Local Similarity	60.4%; Pred. No. 2.8e-23;	
Matches 191; Conservative	0; Mismatches 125; Indels 0; Gaps 0;	

9 TCCCGGCGCCGTCGCGCAGATCTCTCGATGCGCACAGATCTTAAATTCAAAAACCTGG

68

Db 1650 TCCACCCCTCCATCCGCTTACTGATGGCGCAATGCTTCAAGTCGACGATATCG 1591
 QY 69 CCGTTGAGGCAAAAAGATCTTGGGTAAATGCTGTCCTCAATTCATGCGCACCT 128
 Db 1590 GGCTGCGGGCGCAAAAGATCTCTCACGCTATGGCCAAAAAATGCGCCAAAGACT 1531
 QY 129 TAATAGACGACACAAAGTAGAGGTGCTGATGAGCTCTTACAGAGTACCGAGGAGTACA 188
 Db 1530 TCATCATGGCAACGCGCTGCTGTCGACCACTCTTACAGGCTCTGCAAGATGACCA 1471
 QY 189 CCGGAAACAAAGAGGAGGAGAGATCATCAAGAACTCATCAAGACATCAAGC 248
 Db 1470 CCGGAAACAAAGGCGGAGAGAGCTGATCAAGAACTCATCAAGATGATGATCAAGA 1411
 QY 249 TGGCCATCTTTATAGAAATATATAGTTTAAATCAAGATGAGCTTATGAGAAAT 308
 Db 1410 TCGGTGTGCTCCACCGGAAACATCATGTCAGCGACGAGAGCTCAGAGGCGAGCTCT 1351
 QY 309 TTAAGAGAAAGTTCA 324
 Db 1350 TTAAGAGAAAGTTCA 1335

RESULT 39

AA192435/c
 ID AA192435 standard; cDNA; 400 BP.

AA192435;

06-NOV-2001 (first entry)

Human polynucleotide SEQ ID NO 12495.

Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 tissue growth factor; immunomodulatory; cancer; leukaemia;
 nervous system disorders; arthritis; inflammation; ss.

Homo sapiens.

MO200164835-A2.

07-SEP-2001.

26-FEB-2001; 2001MO-US004927.

28-FEB-2000; 2000US-00515126.

18-MAY-2000; 2000US-00577409.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-514838/56.

P-PSDB; AA012504.

Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 and treating e.g. leukemia, inflammation and immune disorders.

Claim 1; SEQ ID NO 12495; 1399pp + Sequence listing; English.

The invention relates to human polynucleotides (AA179941-AA193841) and
 the encoded proteins (AA000010-AA013910) that exhibit activity relating to
 cytokine, cell proliferation or cell differentiation or which may induce
 production of other cytokines in other cell populations. The
 polynucleotides and polypeptides are useful in gene therapy, vaccines or
 peptide therapy. The polypeptides have various cytokine-like activities,
 e.g. stem cell growth factor activity, haematopoiesis regulating
 activity, tissue growth factor activity, immunomodulatory activity and
 activity/inhibit activity and may be useful in the diagnosis and/or
 treatment of cancer, leukaemia, nervous system disorders, arthritis and
 inflammation. Note: The sequence data for this patent did not form part
 of the printed specification, but was obtained in electronic format

CC directly from WPI at ftp.wipo.int/pub/published_pcr_sequences
 XX Sequence 400 BP; 80 A; 122 C; 114 G; 84 T; 0 U; 0 Other;

Query Match 16.64; Score 99.8; DB 4; Length 400;
 Best Local Similarity 68.84; Pred. No. 5.4e-19;
 Matches 137; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 23 CCGACTCTCCGATGCGCACAGATGCTTAAATTCAAAAACCTGCGCTTACGGCACAA 82
 Db 199 CTGTGTCCTCCGAGGCGCATGACACCTTACGACCAAGAGCTGCTGACAGGCGCAG 140
 QY 83 AAGAAGATCTTGGGTAAATGCTGTCCTCAATTCATGCCACCACTTAATAGACACACA 142
 Db 139 AAGAAAGCTCTGAGTAAAGATGAGGCTCCAGGCGAGTGCTGCGCTGTGATGACACC 80
 QY 143 AGTAGAGGTGCTGATGAGCTCTTACAGAGTACCAAGGAGTACCCCAAAACAAAG 202
 Db 79 AGCAGTAGAGGTGCTGATGAGCTGTACCGCCACCAAGGAGTTACGCGCAGCGCAG 20
 QY 203 GAGGACAGAAAGATCATCA 221
 Db 19 GAGGCCAAGAAAGATGCTCA 1

RESULT 40

AAH04721
 ID AAH04721 standard; cDNA; 714 BP.

AAH04721;

26-JUN-2001 (first entry)

Human cDNA clone (5'-primer) SEQ ID NO:1556.

Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-00116126.

29-JUL-1999; 99JP-00248036.

27-AUG-1999; 99JP-00300253.

11-JAN-2000; 2000JP-00118776.

02-MAY-2000; 2000JP-00183767.

09-JUN-2000; 2000JP-00241899.

(HELI-) HELIX RES INST.

Ota T, Isegai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J,
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI; 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 length cDNAs defined in the specification, and for the detection and/or
 diagnosis of the abnormality of the proteins encoded by the full-length
 cDNAs.

Claim 1; SEQ ID NO 1556; 2537pp + Sequence listing; English.

The present invention describes primer sets for synthesizing 5602 full-
 length cDNAs defined in the specification. Where a primer set comprises:
 (a) an oligo-dT primer and an oligonucleotide complementary to the
 complementary strand of a polynucleotide which comprises one of the 5602
 nucleotide sequences defined in the specification; where the
 oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 of an oligonucleotide comprising a sequence complementary to the
 complementary strand of a polynucleotide which comprises a 5'-end

OS Homo sapiens.

XX	PN	XX	WO200210449-A2.
XX	XX	PD	07-FEB-2002.
XX	PF	20-JUL-2001;	2001WO-IB001903.
XX	PR	28-JUL-2000;	2000US-0221607P.
XX	PR	02-MAY-2001;	2001US-0287724P.
XX	PA	(COMP-)	COMPUGEN INC.
XX	PI	Shoshan A, Messerman A, Mintz E, Mintz L, Faigler S;	
XX	PI	WPI, 2002-257383/30.	
XX	PT	New oligonucleotide libraries comprising oligonucleotides which	
XX	PT	selectively hybridize to mRNAs transcribed from a transcription unit of a	
XX	PT	genome, useful for detecting tissue-, pathology-, and developmental-	
XX	PT	specific genes.	
XX	PS	Example 1; SEQ ID NO 23152; 47bp; English.	
XX	CC	The present invention describes oligonucleotide libraries for detecting	
XX	CC	messenger RNAs that populate a (sub-)transcriptome, where the (sub-	
XX	CC)transcriptome comprises messenger RNAs transcribed from multiple	
XX	CC	oligonucleotides, each capable of hybridizing selectively to a set of	
XX	CC	messenger RNAs transcribed from a given transcription unit of the genome,	
XX	CC	which encodes one or more messenger RNA splice variants. The	
XX	CC	oligonucleotide libraries are useful for detecting mRNAs from a	
XX	CC	biological sample, in expression profiling studies, in qualitatively or	
XX	CC	quantitatively characterising the corresponding transcriptome, and in	
XX	CC	detecting RNA transcripts and splice variants of human or animal	
XX	CC	transcriptomes. The libraries may also be used as specialised mini	
XX	CC	libraries to detect transcripts of a sub-transcriptome under a particular	
XX	CC	biological or pathological state, and so allowing the detection of tissue	
XX	CC	- and pathology-specific genes such as those genes only expressed in	
XX	CC	specific tissue under a specific pathological condition; to detect	
XX	CC	developmental specific genes; and to detect RNA transcripts and splice	
XX	CC	variants of a transcriptome of a patient suffering from a particular	
XX	CC	disorder. ABN7255 to ABN59589 represent oligonucleotide sequences from	
XX	CC	rats, humans and mice, which are used in the exemplification of the	
XX	CC	present invention. N.B. The sequence data for this patent did not form	
XX	CC	part of the printed specification, but was obtained in electronic format	
XX	CC	directly from WIPO at http://wipo.int/pub/published_pct_sequences	
XX	XX	Sequence 60 BP; 22 A; 14 C; 14 G; 10 T; 0 U; 0 Other;	
XX	XX	Query Match 10.0%; Score 60; DB 6; Length 60;	
XX	XX	Best Local Similarity 100.0%; Pred. No. 1.3e-07;	
XX	XX	Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
XX	XX	199 GAAGGAGCGAGAGAGATCATCAAGAACTCATCAAGACAGTCATCAAGTGGCCATTCT 258	
XX	XX	1 GAAGGAGCGAGAGAGATCATCAAGAACTCATCAAGACAGTCATCAAGTGGCCATTCT 60	
XX	XX	RESULT 43	
XX	XX	ID ACF25360 standard; DNA; 7095 BP.	
XX	XX	ACF25360;	
XX	XX	AC	
XX	XX	22-SEP-2003 (first entry)	
XX	XX	Rat microtubule associated protein IB gene.	
XX	XX	spinal cord; neuropathic pain; central sensitisation pain; pain;	
XX	XX	analgesic; gene therapy; gene; ss.	
XX	XX	Rattus norvegicus.	
XX	XX		

Pt		EP1284298-A2.
Pd		19-FEB-2003.
Pf		26-JUL-2002; 2002EP-00255229.
Px		
Px		
Pr		27-JUN-2001; 2001GB-00018354. 07-FEB-2002; 2002GB-00002883.
Pa	(WARN)	WARNER LAMBERT CO.
Pi	Brookbank RA,	Dixon AK, Lee K, Pinnock RD;
Dr	WPI; 2003-543489/52.	
Rr	P-PsDB; AEM04816.	
Tp	Use of an isolated gene sequence in the screening of compounds for diagnosing or treating pain.	
Xx	Claim 1; Page 125-127; 189pp; English.	
Fs	The invention relates to a novel isolated gene sequence that is downregulated in the spinal cord of a mammal in response to mechanistically distinct first and second models of neuropathic or central sensitisation pain, useful in the screening of compounds for diagnosing or treating pain. A protein encoded by a gene of the invention has analgesic activity. A polynucleotide of the invention may have a use in gene therapy. The gene sequence is useful for preparing a composition for diagnosing or treating pain. The present sequence represents a gene of the invention	
Cc		
Sq	Sequence 7095 BP; 2124 A; 1856 C; 1799 G; 1316 T; 0 U; 0 Other;	
Qy	Query Match	7.9%; Score 47.4; DB 9; Length 7095; Best Local Similarity 48.7%; Pred.No. 0.008; Matches 129; Conservative 0; Mismatches 136; Indels 0; Gaps 0;
Db	53 AATTCCAAAACTGGCGTTTCAGGCACAAAGAATCTTGGAATAAATGTGTCCAA 112 1645 AAACC CGAAAAGAAAAGAGAGAACGCCAAGAGAGAGTGGCTTA AAAAGAGACA AA 1704	
Oy	113 TC CATCGC CCA CCTTA ATTAG AC GA CA CA AG TAG TG AG GT CTG AG TGA C TT ACA GA 172	
Db	1705 ACTCCCTCTCAAGAAAGACGAGAACCCCAAAAAGMAAGGCGGAAGAGCATCAAGAA 1764	
Oy	173 GTG ACCAGGAGTACACCCA AAAACA GA AGS CAG AAGATCATCA GA ACCTATC 232	
Db	1765 GAAATTCAAAAAGGAAGAAAAGACTGAAGAAAAGGTGAAGAAAGAAACGCCCTG 1824	
Oy	233 AAGACATGATCAAGCTGGCCATCTTTATVAGAAATATCAGTTTAATCAAGTAGACTA 292	
Db	1825 AAGACGCCCAAGAGAGGTGAAGAAAAGACAGAAAGAAAGATTAAAAAGAGAGAGAG 1884	
Oy	293 GCATTGATGAGAAATTTAAGAGA 317	
Db	1885 GAACCCAAAAAGAGATTAAAGAGA 1909	
Result 44		
ID	ACH78513 standard; DNA; 543 BP.	
Ach78513:		
DT	29-JUN-2004 (first entry)	
Hm	Human genome derived single exon probe #11708.	
Kw	Human: probe; ss; gene expression; single exon probe; microarray;	
Km	alternative splicing event; genomic alteration.	
Os	Homo sapiens.	

PN US2003194704-A1.
 XX 16-OCT-2003.
 PD
 XX
 PF 03-APR-2002; 2002US-00029386.
 PR 03-APR-2002; 2002US-00029386.
 XX
 XX (PENN/) PENN S G.
 PA (RANK/) RANK D R.
 PA (HANZ/) HANZEL D K.
 XX
 XX Penn SG, Rank DR, Hanzel DK;
 PI
 DR MPI, 2004-119264/12.
 XX
 XX New human genome-derived single exon nucleic acid probes useful for human
 PT gene expression analysis, for identifying or characterizing alternative
 PT splicing events, for assessing genomic alterations or as tools for
 PT surveying tissues.
 XX
 XX
 PS Claim 15; SEQ ID NO 11708; 80pp; English.
 XX
 XX The invention relates to a nucleic acid probe for measuring human gene
 CC expression, comprising any of the 27,400 fully defined nucleotide
 CC sequences in the specification, or their complements or fragments, and
 CC encoding at least 8 amino acids of any of the 688 amino acid sequences
 CC fully defined in the specification. The probe is a single exon probe that
 CC hybridizes under high stringency conditions to a nucleic acid molecule
 CC expressed in human cells or tissues. Also included are a spatially-
 CC addressable set of single exon nucleic acid probes for measuring human
 CC gene expression (comprising a plurality of single exon nucleic acid
 CC probes cited above, where each of the plurality of probes is separately
 CC and addressably isolatable or amplifiable from the plurality), a single
 CC exon microarray for measuring human gene expression, a method of
 CC measuring human gene expression, a vector comprising the single exon
 CC probe cited above, an ORF-encoded peptide comprising at least 8
 CC contiguous amino acids of any of the above-mentioned amino acid
 CC sequences (optionally with conservative amino acid substitutions), an
 CC isolated antibody that binds specifically to a peptide cited above,
 CC methods of selling and/or licensing single exon probes or microarrays to
 CC a customer desiring to measure gene expression, a method of providing
 CC human gene expression data by subscription, and a computer-readable
 CC storage medium which contains a database having a plurality of records
 CC (each record including data on the expression of a single exon probe
 CC cited above. The probe, methods and apparatus are useful in gene
 CC expression analysis. The probes may be used as tools for surveying
 CC tissues to detect the presence of expressed messages that contain their
 CC specific exon, or in constructing genome-derived single exon microarrays.
 CC In addition, the probes are used in identifying and characterizing
 CC alternative splicing events, in detecting and characterizing gross
 CC alterations in the genomic locus that includes their exon, in assessing
 CC smaller genomic alterations, in priming the synthesis of nucleic acids,
 CC or in expressing the ORF-encoded peptide. The present sequence is a human
 CC single exon probe of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docid=20030194704
 XX
 XX
 SQ Sequence 543 BP; 137 A; 128 C; 119 G; 159 T; 0 U; 0 Other;
 Query Match 7.8%; Score 47; DB 12; Length 543;
 Best Local Similarity 58.0%; Pred. No. 0.0023;
 Matches 83; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
 QY 459 ATGAGCGGTTAATAATGCTTGTGATCATTTTTCAGATTGTGAATTTTGGTCGCTTGT 518
 DB 1 ACGGGGCGATCAACGAGCTTTTAACCATTTGCCGATGGAGTTCCTCCACCCCTCT 60
 QY 519 AATATCTTTTGGGAATTTTAAACCCACTTACAAAACATATGTATTAACAAA 578
 DB 61 ATATGCTGATGAGAGCTGTAGGCCCAACCTCAAGAGGATTTGTGAAGAAATCAATTAAGT 120

QY 579 TGTGTGATGAAGAGAACATATGA 601
 DB 121 TGTGTGATGAAGAGAACATCTTTTA 143
 RESULT 45
 ID ABQ76615/c
 ABQ76615 standard; cDNA; 15790 BP.
 XX
 AC ABQ76615;
 XX
 DT 21-NOV-2002 (first entry)
 XX
 DE C. albicans BAX-associated cDNA fragment SEQ ID 655.
 XX
 KW Bax; Bax-resistance; cytosolic; fungicide; immunosuppressive; virucide;
 KW vasotropic; vaccine; gene therapy; proliferative disorder; cancer;
 KW apoptosis; fungal; yeast; infection; autoimmune disease; ischaemia;
 KW neurodegeneration; cell death; ss.
 XX
 OS Candida albicans.
 XX
 PN WO200264766-A2.
 XX
 PD 22-AUG-2002.
 XX
 PF 21-DEC-2001; 2001WO-BP015398.
 XX
 PR 22-DEC-2000; 2000EP-00870318.
 PR 04-JAN-2001; 2001EP-00870002.
 PR 09-JAN-2001; 2001EP-00870003.
 XX
 PA (JANC) JANSEN PHARM NV.
 XX
 PI Contreras RH, Eberhardt I, Luyten WHML, Reekmans RJ;
 XX
 DR MPI; 2002-667002/71.
 DR P-PSDB; ABG93349.
 XX
 PT New isolated nucleic acid representing a synthetic BAX-gene, useful as
 PT medicament for treating, preventing and/or alleviating yeast or fungal
 PT infections or proliferative disorders, or for preventing apoptosis in
 PT certain diseases.
 XX
 PS Claim 36; Fig 2; 344pp; English.
 XX
 XX This invention describes a novel nucleic acid representing a synthetic
 CC Bax gene. The Bax gene of the invention is useful for identifying Bax-
 CC resistant yeast or fungi, identifying, or obtaining and identifying
 CC Candida spp. sequences that are differentially expressed in a pathway
 CC eventually leading to programmed cell death or identifying inhibitors or
 CC inhibitor sequences of Bax-induced cell death. The products of the
 CC invention have cytostatic, fungicide; immunosuppressive, virucide and
 CC vasotropic activity and can be used in vaccines or for gene therapy. The
 CC isolated nucleic acids, polypeptides, pharmaceutical compositions,
 CC antisense molecules and antibodies are useful as medicaments or in
 CC preparing a medicament for treating, preventing and/or alleviating
 CC diseases associated with yeast or fungi or proliferative disorders, such
 CC as cancer, or for preventing apoptosis in certain diseases. The compounds
 CC or polypeptides, or the genetically modified organism are useful for
 CC preparing a medicament for modifying the endogenous flora of humans and
 CC other mammals. The vaccine is useful for immunising against yeast or
 CC fungal infections. Apoptosis-related diseases include autoimmune disease,
 CC ischaemia, diseases related with viral infections or neurodegenerations.
 CC This sequence represents a polynucleotide associated with the Bax gene
 CC described in the disclosure of the invention
 XX
 XX
 SQ Sequence 15790 BP; 5181 A; 2570 C; 2199 G; 5840 T; 0 U; 0 Other;
 Query Match 7.1%; Score 42.4; DB 6; Length 15790;
 Best Local Similarity 55.4%; Pred. No. 0.36;
 Matches 82; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

ALIGNMENTS

```
RESULT 1
US-09-799-451-537
; Sequence 537, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Kyle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: pf_fl_genes Version 2.0
; SEQ ID NO 537
; LENGTH: 1175
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (113)..(664)
US-09-799-451-537

Query Match          30.3%; Score 182; DB 4; Length 1175;
Best Local Similarity 58.7%; Pred. No. 6.1e-47;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGATGGCCACAGATGCTTTAATTCAGAAAACCTGSCCTTCAGGACAGAAAGAGA 89
DB 102 CTCAGAGACCCATGGAGTCCTTCAGCTCAAGAGCCTGGCACTGCAAGCAGAGAGAAC 161
QY 90 TCTTGGGTAAATGCTGTCCAAATCATGCGCACCTTAAATAGCGACAAAGTAGTG 149
DB 162 TACTGATGATAGTGGGGGTGCGCTGTGTCATCTTCAATAGATGAGACAAGCATG 221
QY 150 AGGTGCTGATGAGCTCTTACAGAGTACACAGGAGTAGACCCAGAAACAGAGAGGAG 209
DB 222 AGGTGCTGATGAGCTCTTACAGAGTACACAGGAGTAGACCCAGAAACAGAGAGG 281
QY 210 AGAAGATCATCAAGAACCTCATCAAGACAGTATCAAGCTGGCCATCTTTATAGGATA 269
DB 282 AGCGCGTATCAAGGACCTGATCAAGAGTGCCATCAAGGTGGCTGTGCTGACCCGAA 341
QY 270 ATCAGTTTATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTATCAGC 329
DB 342 GCTCCTTGGCCCAAGTAGCTGGCCCTGCTAACCCGCTTTCGCCAGAAAGCTGGCAG 401
QY 330 TTGCTATGACCGGCTGAGTTTCATCAGTGGATTATACCTTTGACCGGAATGTTAT 389
DB 402 GTGCGATGACGGCATTTACTTGTGTGAGGTAGATTCTTACCTTTCGAGGCTGCTGTT 461
QY 390 CCAGGCTGTTAATGATGACAGAGAGTCTGACCAACCAATTCATTGAGCCGACCTCA 449
```

```
DB 462 CTGACCTGTGACCGAGTGGCCGAGTGTGCTAGAGTGGTGGAAACCACTACGCC 521
QY 450 CCAAGTCACATGAGACGGGTATATATGCTTTGATCATTTTTCAGATGTGAAATTTGG 509
DB 522 CCAAGTCACATGAGCCGATCCGACAGTGTGGATCATCTTCTTACCCAGGTGCTCA 581
QY 510 CTGACCTGTATATCTTTTGGGAATTTTAAACCCACTTACAAAAACTATGTAGTGA 569
DB 582 CGGCCCCCTTATGGGCTT-----GACTTCATCTACAGACCTTGGCAAGATCTGTAC 635
QY 570 TCACAAATATGTTGATGAGAGAGAACATATGA 601
DB 636 TCAGAGAGCTGTACACGAGAGGAGCTCTGA 667

RESULT 2
US-09-513-999C-14427
; Sequence 14427, Application US/0951399C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Meline Edwards, J.B.
; APPLICANT: Duclet, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14427
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 139
; OTHER INFORMATION: w=a or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 166
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 170
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 208
; OTHER INFORMATION: k=g or t
; NAME/KEY: misc_feature
; LOCATION: 221
; OTHER INFORMATION: r=a or g
US-09-513-999C-14427

Query Match          26.5%; Score 159.2; DB 4; Length 340;
Best Local Similarity 98.8%; Pred. No. 4.7e-40;
Matches 158; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 45 ATGCTCTTATTCAGAAAACCTGCGCTTCAGGACAGAAAGAAATCTTGGTAAATGG 104
DB 181 ATGCTCTTATTCAGAAAACCTGCGCTTCAGGACAGAAAGAAATCTTGGGTAATGG 240
QY 105 TGTCCAAATCCATGGCCCACTTAAATAGACGACAAAGTAGAGGTGCTGATGAGC 164
DB 241 TGTCCAAATCCATGGCCCACTTAAATAGACGACAAAGTAGAGGTGCTGATGAGC 300
QY 165 TCTACAGATGACACGAGAGTAGACCCAGAAACAGAGAGGA 204
```


LENGTH: 2754
TYPE: DNA
ORGANISM: Candida albicans
FEATURE:
NAME/KEY: unsure
LOCATION: (50)
OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unknown
US-09-248-796A-4857

Query Match
Best Local Similarity 7.1%; Score 42.4; DB 4; Length 2754;
Pred. No. 0.014;
Matches 82; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 181 GGAGTACACCCAAACAGAGAGAGAGAGATCATCAAGACCTCATCAAGACAGT 240
DB 1072 GAATCAGATCTATATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1131
QY 241 CATCAAGCTGGCCATCTTTATAGATATATCAAGATTAATCAAGATGAGCTAGCATTTGAT 300
DB 1132 AAAAATTTGAAATTTCTGAAGTGAATCAAAATGTTGAATCTCAAGTATCAAAAAGAA 1191
QY 301 GGAGAAATTTAGAGAAAGTTCATCAG 328
DB 1192 TCTGAATTTAAGAGAAAGTACAGAG 1219

RESULT 6

US-09-489-039A-1361
Sequence 1361, Application US/09489039A
Patent No. 6610836
GENERAL INFORMATION:
APPLICANT: Gary Breton et. al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
FILE REFERENCE: 2/09/2004001
CURRENT APPLICATION NUMBER: US/09/489,039A
CURRENT FILING DATE: 2000-01-27
PRIOR APPLICATION NUMBER: US 60/117,747
PRIOR FILING DATE: 1999-01-29
NUMBER OF SEQ ID NOS: 14342
SEQ ID NO 1361
LENGTH: 804
TYPE: DNA
ORGANISM: Klebsiella pneumoniae
US-09-489-039A-1361

Query Match
Best Local Similarity 6.6%; Score 39.4; DB 4; Length 804;
Pred. No. 0.058;
Matches 115; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 123 CCACCTTATATAGACGACACAGTATGAGTGTGATGAGCTCTACAGAGTGAACGAGG 182
DB 434 CAACCTTTAAAGGCAACCATCGAAACCTTGTGTGCTGAGCGCGGGATCAACG 493
QY 183 AGTACACCCAAACAGAGAGAGAGAGATCATCAAGACCTCATCAAGACAGTCA 242
DB 494 CCGCAGCCGCAACAAAGAGTGGCCAAAGAGTCTCGAAAACTACCTGATGACCGATC 553
QY 243 TCAAGCTGGCCATCTTTATAGATATATCAAGATTAATCAAGATGAGTCAAGATTCAGT 302
DB 554 AGGCTCTGAGACCGCTTAACAACGACCAACCGCTGGGTCCCGTGGCGCTGAATCTTCC 613
QY 303 AGAAATTTAAGAAAGTTCATCAAGCTGTATGACCTGTGCACTTTCCATCAAGTGG 362
DB 614 AGGAAATCTGAAAGATCCGCGCATCGTCCCAACATGCGGAAGCCCAAGAAAGGCG 673
QY 363 A 363
DB 674 A 674

RESULT 7
US-09-436-699C-19

Sequence 19, Application US/09436699C
Patent No. 6696619
GENERAL INFORMATION:
APPLICANT: Farnodu, Omolayo O.
TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetases
FILE REFERENCE: B81266 US NA
CURRENT APPLICATION NUMBER: US/09/436,699C
CURRENT FILING DATE: 2002-09-30
PRIOR APPLICATION NUMBER: 60/107,276
PRIOR FILING DATE: 1998-11-05
SOFTWARE: Microsoft Office 97
SEQ ID NO 19
LENGTH: 1719
TYPE: DNA
ORGANISM: Triticum aestivum
US-09-436-699C-19

Query Match
Best Local Similarity 6.3%; Score 37.8; DB 4; Length 1719;
Pred. No. 0.29;
Matches 87; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 149 GAGTGTCTGATGAGCTCTTACAGAGTGAACAGGAGTACACCCAAACAGAGAGGCA 208
DB 229 GAGCTGTGACAGAGATCATGCTCTCGACGAGGCGGTGGCCAGAGGAGTTCAGGCTC 288
QY 209 GAGAAATCATCAACAACTCATCAAGACAGTATCAAGCTGCGCATTTTATAGAAAT 268
DB 289 GACAAAGATCCGCGACAGAGCTCAACAAACAGCAAGAGATCGCAAGCTCAAGGCCAA 348
QY 269 AATCAGTTAATCAAGATGAGCTAGCATTTGAGAGAAATTTAAGAGA 317
DB 349 AACGAGATCGACGAGAGCTGATACAGAGCAGGAGAGATTTAAGAGA 397

RESULT 8

US-08-916-421B-1/c
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Bult et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
Patent No. 6503729
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
CURRENT FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
PRIOR FILING DATE: 1996-08-22
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Patentin version 3.1
SEQ ID NO 1
LENGTH: 1664976
TYPE: DNA
ORGANISM: Methanococcus jannaschii

FEATURE:
NAME/KEY: misc_feature
LOCATION: (28222)..(28222)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (28257)..(28258)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84773)..(84773)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84808)..(84808)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (84812)..(84812)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature


```
LOCATION: (98120)..(98120)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98159)..(98159)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98239)..(98239)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98266)..(98266)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (98343)..(98343)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (10398)..(10398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (148948)..(148948)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (163385)..(163385)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191989)..(191989)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (191995)..(191995)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (231980)..(231980)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234187)..(234187)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234220)..(234220)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (234814)..(234814)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309398)..(309398)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (309418)..(309418)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312837)..(312837)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (312993)..(312993)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (319226)..(319226)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
```

```
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349473)..(1349473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

```
Query Match 6.1%; Score 36.8; DB 4; Length 1664976;
Best Local Similarity 48.1%; Pred. No. 37;
Matches 104; Conservative 0; Mismatches 112; Indels 0; Gaps 0;
```

```

OY      216  TCATCAGAACCTCATCAAGCAGCATCAAGCTGGCCATCTTATATAGAAATATACGT 275
Db      1399648  TCTTTAATTAATTCCTCATCATATAAAGTCAGTTTCTTTCTTTGTTATATAGCTTCAATTCCTT 1399589
OY      276  TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGTTTCATCAGCTTGCTA 335
Db      1399588  ATTTTTCATATCTTAATCTTATATATATGCCATCTTACAGAGAAATTTCAATTAATCTTGTTA 1399529
OY      336  TGACCGTGTGATTTCCATCAGGTGATTTATACCTTTGACCGGAATGTGTATCCAGGC 395
Db      1399528  TAACTTTTTTACACTCTCCAAATTTTCTTAATATCTAATGACTTAAGTTCAATTTTTGAGA 1399469
OY      396  TGTAAATGAATGACAGAGATGCTGCACCAATCA 431
Db      1399468  TTTTAAATATTTCTGCCTTATCCAGCACTTGCTCA 1399433

RESULT 9
US-09-692-570-1/c
; Sequence 1, Application US/09692570
; Patent No. 6797466
; GENERAL INFORMATION:
; APPLICANT: Buit et al.
; TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus
; Patent No. 6797466
; TITLE OF INVENTION: jannaschii
; FILE REFERENCE: PB275C1
; CURRENT APPLICATION NUMBER: US/09/692,570
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/024,428
; PRIOR FILING DATE: 1996-08-22
; PRIOR APPLICATION NUMBER: US 08/916,421
; PRIOR FILING DATE: 1997-08-22
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1664976
; TYPE: DNA
; ORGANISM: Methanococcus jannaschii
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28252)..(28222)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (28257)..(28258)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84773)..(84773)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84808)..(84808)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (84812)..(84812)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98120)..(98120)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98159)..(98159)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (98239)..(98239)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc feature

```

LOCATION: (98266) .. (98266)	a, t, c, or g
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (98343) .. (98343)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (103598) .. (103598)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (148948) .. (148948)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (163385) .. (163385)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (191989) .. (191989)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (231580) .. (231580)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (234187) .. (234187)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (234220) .. (234220)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (234814) .. (234814)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (309398) .. (309398)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (309418) .. (309418)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (312837) .. (312837)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (312993) .. (312993)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (319226) .. (319226)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (559167) .. (559167)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (559241) .. (559241)	
OTHER INFORMATION: n equals a, t, c, or g	
FEATURE:	
NAME/KEY: misc_feature	
LOCATION: (600992) .. (600992)	

FEATURE:	OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY:	misc_feature
LOCATION:	(622708) .. (622708)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(657081) .. (657081)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(657203) .. (657203)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(674435) .. (674435)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(682442) .. (682442)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(713652) .. (713652)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(741684) .. (741684)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(779455) .. (779455)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(779676) .. (779676)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(855539) .. (855539)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(871619) .. (871619)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(1084930) .. (1084930)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(1096846) .. (1096846)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(1119881) .. (1119881)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(1130881) .. (1130881)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(1310988) .. (1310988)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(1313224) .. (1313224)
OTHER INFORMATION:	n equals a, t, c, or g
FEATURE:	
NAME/KEY:	misc_feature
LOCATION:	(1349473) .. (1349473)
OTHER INFORMATION:	n equals a, t, c, or g

	Query Match	Best Local Match	Similarity	4.1%;	Score 36.8;	DB 4;	Length 1664976;
	Matches	104;	Conservative	0;	Mismatches 112;	Indels 0;	Gaps 0;
Qy	216	TCATCAAGAACCTTCATCAAGACAGCTGCATCAAGCTGGCCATCTTTATATAGAAATATACGT	275				
Db	1399648	TCCTTAATATTCCTCATCAATAAAGTCAGTTTTCCTTTGGTTTATATAGCTTCATCTCTT	1399589				
Qy	276	TTATCAAGATGAGCTGCACTGATGATGAGAAATTTAGAAAGATTCATCAGCTGCTCA	335				
Db	1399588	ATTTTTTCATATTTAAATCTTATATATATGACCATCTTTACGAGAAATTTCAATTAACHTGTTA	1399529				
Qy	336	TGACCGTGGTCAGTTTCCATCAAGTGATATATACCTTTGACCGGAATGTGTATCCAGC	395				
Db	1399588	TAACTTTTTTACACTCTCCAAATTTTCTTAATATCTATGACTTAAGTTCATTTTTTGAGA	1399469				
Qy	396	TGTTAATGAATGACGAGAGATGCTGACCAATCA	431				
Db	1399468	TTTTTAATAATTTTGCTTATCCAGCACTTGCTCA	1399433				

```

RESULT 10
US-09-621-976-2813
; Sequence 2813, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J. Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ. ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 2813
; LENGTH: 832
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 235..399
; US-09-621-976-2813

```

Query Match	6.1%;	Score 36.6;	DB 4;	Length 832;
Best Local Similarity	11.6%;	Pred. No. 0.46;		
Matches	42;	Conservative 165;	Mismatches 154;	Indels 2;
				Gaps 1;
Qy	241	CATCAAGCTGGCCATCTTTATAGGAATTAATCATGTTAATCAAGATGAGCTAGCAT	300	
Db	1	YRWYTWKTTTYYACWTWKMKMSWYMMWTKYTKTWTWRMRKKKKKMKYKWTWTWY	60	
Qy	301	GGAGAAATTTAAGAAGAAAGTTTCATCAGCTGTGATAGACCGTGTGATTTCCATCAGT	360	
Db	61	RYAMWGYKKKAMCERTKTKKKKKKKGWWMMYWGWRBSYMAWTRTWTWGYAYRSMYWMR	120	
Qy	361	GGATTATACTTTGACCGGAAATG--TTATCCAGCGCTTTAAATGAATGACAGAGATG	418	
Db	121	YRCMKKAYYRKTCYSSBKWTWWRKMKATTTWMMKTYMAATRYMMWCWTKRMS	180	
Qy	419	CTGACCAAAATCATTCAGCGCACCTCAGCCCAAGTACACATGAGCGGTTAATATGTC	478	
Db	181	IMYCMWMMGKARKMSWTMRKRSYASABAKKCTCYSSGMAWSKTMKMRMRPRGHAITGAM	240	
Qy	479	TTTGATCATTTTTCAGATGTGAAATTTTGGCTGCTGTATATAATCTTTTGGGAAATTT	538	
Db	241	KAMRASCMERKRYAGRKSTYSKMMWCMTRSMKCYTKABWTGYCYRKGGMMKGRMY	300	
Qy	539	AAACCCCACTTACAAAATATGATGAGATGATCAACAAAATGTTGATGAAGAACAATA	598	
Db	301	ASKKYMKRMMWCARARYSTGRLASMMBRMYTMMKMKYKAMAPLARMAMMMWAW	360	


```
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7337
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-7337

Query Match
Best Local Similarity 51.6%; Score 36.2; DB 4; Length 502;
Pred. No. 0.45;
Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 386 TTATCCAGCGCTGTAATGATGACAGAGATGCTGCACCAATCATTCAGCCGCACTC 445
DB 398 TTGACCATTAATTAATCAATTCATTCGACATTAATGACGTTGCAACAAAG 339
QY 446 ACTGCCAAGTCACATGACGCGTTAATATATGCTTTGATCATTTTTCAGATTGGAATT 505
DB 338 CCAAGTAACCCACAGATATGTTGTTGTCATTTTGTATTTTGGTTGCTTGTTG 279
QY 506 TTGGCTGCTTGTATATCTCTTTTGGGAATTTAAACCCA 546
DB 278 GTTGATTAATCTGTAATATACGTTTGTGACGCTGAACCGCA 238

RESULT 15
US-09-270-767-22619/c
; Sequence 22619, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 22619
; LENGTH: 502
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-22619

Query Match
Best Local Similarity 51.6%; Score 36.2; DB 4; Length 502;
Pred. No. 0.45;
Matches 83; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 386 TTATCCAGCGCTGTAATGATGACAGAGATGCTGCACCAATCATTCAGCCGCACTC 445
DB 398 TTGACCATTAATTAATCAATTCATTCGACATTAATGACGTTGCAACAAAG 339
QY 446 ACTGCCAAGTCACATGACGCGTTAATATATGCTTTGATCATTTTTCAGATTGGAATT 505
DB 338 CCAAGTAACCCACAGATATGTTGTTGTCATTTTGTATTTTGGTTGCTTGTTG 279
QY 506 TTGGCTGCTTGTATATCTCTTTTGGGAATTTAAACCCA 546
DB 278 GTTGATTAATCTGTAATATACGTTTGTGACGCTGAACCGCA 238

RESULT 16
US-09-949-016-23836/c
; Sequence 23836, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23836
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-23836

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 601;
Pred. No. 0.5;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 184 GTACACCCAAAACAAGAGAGGACAGAAATCATCATCAAGCTCATCAAGACATCAT 243
DB 202 GAACACACCATTAATAAAGTAGGCAAGATATTAACAGTCTTCAAAAGAACAT 143
QY 244 CAAGTGGCCATTCTTATAGAAATATCATGTTAATCAAGATGACATGATGATGCA 303
DB 142 TTATGCGGCAACAACATATGAGAAAAGCTCAACATCGCTGATCTTAGAGAAATT 83
QY 304 GAAATTTAAGAAAGAAAGTTCATCATCTGCTAT 336
DB 82 TATACCTATGACCCCAATTCACCAATCATAT 50

RESULT 17
US-09-949-016-173088/c
; Sequence 173088, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 173088
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-173088

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 601;
Pred. No. 0.5;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
```

```
QY 184 GTACACCCAAAACAAGAGAGGACAGAAATCATCATCAAGCTCATCAAGACATCAT 243
DB 202 GAACACACCATTAATAAAGTAGGCAAGATATTAACAGTCTTCAAAAGAACAT 143
QY 244 CAAGTGGCCATTCTTATAGAAATATCATGTTAATCAAGATGACATGATGATGCA 303
DB 142 TTATGCGGCAACAACATATGAGAAAAGCTCAACATCGCTGATCTTAGAGAAATT 83
QY 304 GAAATTTAAGAAAGAAAGTTCATCATCTGCTAT 336
```

DB 82 TATACCTATGACCCCAATTCACACATCATAT 50

RESULT 18
US-09-540-236-1642/c
; Sequence 1642, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; PRIORITY FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 1642
; LENGTH: 630
; TYPE: DNA
; ORGANISM: M. catarrhalis
US-09-540-236-1642

Query Match
Best Local Similarity 60.8%; Score 36.2; DB 4; Length 630;
Pred. No. 0.52;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 281 CAAGATGACCTAGCATTTGAGAAATTTAAGAAAGTTGATGCTTGTATGACC 340
DB 439 CTACATGTCGACGATTTGCTGGACATCTCATTAAGTCGCTGATGCGACACCTA 380
QY 341 GTGGTCAGTTTCCATCAGGTGATTTACCTTTGACC 377
DB 379 ATTAATAATTGGCTCTAGTGTATCATGCCCTTGATC 343

RESULT 19
US-09-596-002-15
; Sequence 15, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT: Lagace, Robert, E.
; APPLICANT: Paterson, Chandra
; APPLICANT: Berg, Kim, L.
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE: PM-0008-4 US
; CURRENT APPLICATION NUMBER: US/09/596,002
; PRIORITY FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/140,121
; PRIOR FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 28626
; TYPE: DNA
; ORGANISM: M. catarrhalis
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte template ID No. 6632636 15
; PUBLICATION INFORMATION:
US-09-596-002-15

Query Match
Best Local Similarity 6.0%; Score 36.2; DB 4; Length 28626;
Pred. No. 5.1;
Matches 59; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 281 CAAGATGACCTAGCATTTGAGAAATTTAAGAAAGTTGATGCTTGTATGACC 340
DB 17972 CTACATGTCGACGATTTGCTGGACATCTCATTAAGTCGCTGATGCGACACCTA 18031
QY 341 GTGGTCAGTTTCCATCAGGTGATTTACCTTTGACC 377
DB 18032 ATTAATAATTGGCTCTAGTGTATCATGCCCTTGATC 18068

RESULT 20
US-09-949-016-16647
; Sequence 16647, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16647
; LENGTH: 32039
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16647

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 32039;
Pred. No. 5.4;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 184 GTACACCCCAAAAGAGAGAGAGAGATCATCAAGAACTCATCAAGACATCAT 243
DB 17364 GAAACAACACCATCAAAAGTAGGCAAGATTTAAAGCTCTCTCAAAAGAGAT 17423
QY 244 CAAGCTGGCCATCTTTATAGAAATATATCATGTTATATCAAGATGACATTTGATGA 303
DB 17424 TTATGTGGCCCAAAACATATATAGAGAAAAGCTCAACATCGCTGATCTTATAGAGAAATTT 17483
QY 304 GAAATTTAAGAAAGATTCATCAGCTTGCTAT 336
DB 17484 TATACCTATGACCCCAATTCACACCATCATAT 17516

RESULT 21
US-09-949-016-12028
; Sequence 12028, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIORITY FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12028
; LENGTH: 32043
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12028

Query Match
Best Local Similarity 52.3%; Score 36.2; DB 4; Length 32043;
Pred. No. 5.4;
Matches 80; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 184 GTACACCCCAAAAGAGAGAGAGAGATCATCAAGAACTCATCAAGACATCAT 243


```
/ Patent No. 6399297
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Brenda F.
/ APPLICANT: Comsert, Lex M.
/ APPLICANT: Monia, Bret P.
/ APPLICANT: Xu, Xiaoxing S.
/ TITLE OF INVENTION: ANTISENSE MODULATION OF TRAF EXPRESSION
/ FILE REFERENCE: ISPH-0321
/ CURRENT APPLICATION NUMBER: US/09/167,109
/ PENDING FILING DATE: 1998-10-06
/ NUMBER OF SEQ ID NOS: 228
/ SEQ ID NO 3
/ LENGTH: 2455
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (218)..(1924)
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: UZ1092
/ DATABASE ENTRY DATE: 1995-03-23
US-09-167-109-3
```

```
Query Match
Best Local Similarity 51.6%; Score 35.4; DB 3; Length 2455;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
QY 72 TTCAGGACCAAAAGAGATCTTGGGTAAATGTGTCAAAATCCATCCGACACCTTAA 131
DB 1122 TTGAGGACCAAAAGAAATGCTTCGAAATATGATTAATCCATTTTACAGCGAG 1181
QY 132 TAGACGACACAAGTAGTGCTGTGATGAGCTCTACAGAGTACAGGAGTACACCC 191
DB 1182 TGATGACAGCCAGCAGAGGAAATCTGAAGAGCTTGACAGAGGATCCGGCCCTCCGGC 1241
QY 192 AAAACAGAGGAGGAGAGAGATCATCAAGACCT 228
DB 1242 AGAACTGGAGAGAGCAGACAGCATGAGAGCAGCGT 1278
```

```
RESULT 26
US-09-949-016-468
/ Sequence 468, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PENDING FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 468
/ LENGTH: 2455
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-468
```

```
Query Match
Best Local Similarity 51.6%; Score 35.4; DB 4; Length 2455;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
QY 72 TTCAGGACCAAAAGAGATCTTGGGTAAATGTGTCAAAATCCATCCGACACCTTAA 131
DB 1122 TTGAGGACCAAAAGAAATGCTTCGAAATATGATTAATCCATTTTACAGCGAG 1181
```

```
QY 132 TAGACGACACAAGTAGTGCTGTGATGAGCTCTACAGAGTACAGGAGTACACCC 191
DB 1182 TGATGACAGCCAGCAGAGGAAATCTGAAGAGCTTGACAGAGGATCCGGCCCTCCGGC 1241
QY 192 AAAACAGAGGAGGAGAGAGATCATCAAGACCT 228
DB 1242 AGAACTGGAGAGAGCAGACAGCATGAGAGCAGCGT 1278
```

```
RESULT 27
US-08-813-323C-4
/ Sequence 4, Application US/08813323C
/ Patent No. 6822070
/ GENERAL INFORMATION:
/ APPLICANT: Baltimore, David
/ APPLICANT: Cheng, Genhong
/ APPLICANT: Ye, Zheng-sheng
/ APPLICANT: Lederman, Seth
/ APPLICANT: Cleary, Aileen
/ TITLE OF INVENTION: Truncated Craf-1 Inhibits CD40 Signalling
/ FILE REFERENCE: 0575/50659
/ CURRENT APPLICATION NUMBER: US/08/813,323C
/ PENDING FILING DATE: 1997-03-10
/ NUMBER OF SEQ ID NOS: 11
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 4
/ LENGTH: 2455
/ TYPE: DNA
/ ORGANISM: Homo Sapiens
US-08-813-323C-4
```

```
Query Match
Best Local Similarity 51.6%; Score 35.4; DB 4; Length 2455;
Matches 81; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
QY 72 TTCAGGACCAAAAGAGATCTTGGGTAAATGTGTCAAAATCCATCCGACACCTTAA 131
DB 1122 TTGAGGACCAAAAGAAATGCTTCGAAATATGATTAATCCATTTTACAGCGAG 1181
QY 132 TAGACGACACAAGTAGTGCTGTGATGAGCTCTACAGAGTACAGGAGTACACCC 191
DB 1182 TGATGACAGCCAGCAGAGGAAATCTGAAGAGCTTGACAGAGGATCCGGCCCTCCGGC 1241
QY 192 AAAACAGAGGAGGAGAGAGATCATCAAGACCT 228
DB 1242 AGAACTGGAGAGAGCAGACAGCATGAGAGCAGCGT 1278
```

```
RESULT 28
US-09-949-016-4366
/ Sequence 4366, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PENDING FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 4366
/ LENGTH: 7571
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-4366
```

Query Match	5.9%	Score 35.4	DB 4	Length 8146
Best Local Similarity	46.8%	Pred. No. 4.3		
Matches 111	Conservative	0	Mismatches 126	Indels 0
			Gaps	0

QY	80	CAAAAGAGATCTTGGGTAAATGCTGTCCAATCCATGCCACCACTTAAATAGACGAC	139
Db	2145	CCAAAAGAAAGAGTGGCTTAAABAGAGAGCAAAACCTATCAAGAGAGGAAAAACCA	2204
QY	140	ACAACTAGTAGGTCGTCGATGAGCTTACAGAGTGACCGAGGAGTACCCAAAAACAAG	199
Db	2205	AAAAAGGAAGAGGTGAAAAAAGAAAGTCAAAAAAAGAGATCAAGAAAGAAAGAAAAAGAA	2264
QY	200	AAGGAGGAGAGGAAGATCATCAAGAACTCTCAAGACAGTATCAAGCTTGCCATTCTT	259
Db	2265	CCCAAGAAAGAGGTTTAAAGAAAGAAACCCGCCCAAGAAAGTCAAGAAAGAAAGTAAAGAG	2324
QY	260	TATAGGAATTAATCACTTTTAATCAAGATGAGCTAGCGATGAGGAATTTTAAGAG	316
Db	2325	GAAGAGAAAGGAAGACTGAAAAAGAGAGAAAGAAACCAAAAAAGAAATTAAGAG	2381

RESULT 30.
US-09-566-921-32
; Sequence 32, Application US/09566921
; Patent No. 668288

```

RESULT 31
US-09-949-016-16108
; Sequence 16108, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16108
; LENGTH: 101300
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1) _ (101300)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16108

```

Query Match	Score	DB	Length
5.9%	35.4	4	101300

Best Local Similarity 46.8%; Pred. No. 19;
Matches 111; Conservative 0; Mismatches 126; Indels 0; Gaps 0;
QY 80 CAAAGAGATCTGGGTAATAATGTCCTCAATCCATGCCACCTTAATAGAGC 139
DB 89410 CCAAGAGAGATCTGGGTAATAATGTCCTCAATCCATGCCACCTTAATAGAGC 89469
QY 140 ACAAGTAGAGAGTGTGATGAGCTCTACAGAGTACAGGAGAGTACCCAAACAG 199
DB 89470 AAAAGAGAGAGTGTGATGAGCTCTACAGAGTACAGGAGAGTACCCAAACAG 89529
QY 200 AAGAGGACAGAGAGTGTGATGAGCTCTACAGAGTACAGGAGAGTACCCAAACAG 259
DB 89530 CCAAGAGAGAGTGTGATGAGCTCTACAGAGTACAGGAGAGTACCCAAACAG 89589
QY 260 TATAGAGTATATCATGTTTATATCATGAGTACGCTTGTAGAGAAATTTAAGAG 316
DB 89590 GAAAGAGAGAGTGTGATGAGCTCTACAGAGTACGCTTGTAGAGAAATTTAAGAG 89646
RESULT 32
US-09-949-016-17590
Sequence 17590, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17590
LENGTH: 247299
TYPE: DNA
ORGANISM: Human
US-09-949-016-17590
Query Match 5.9%; Score 35.4; DB 4; Length 247299;
Best Local Similarity 56.4%; Pred. No. 33;
Matches 66; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 163 GCTCTACAGAGTACAGGAGGATACCCCAACAGAGAGGAGAGATCATCA 222
DB 180916 GCTCTACATTTACTTAACAAAGAGACTCGAATTTGTGATGAATTAATAAACA 180975
QY 223 GAACCTCATCAAGACGATCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAA 279
DB 180976 GAACCTCATCAAGATCATCAATTCATCTGCCCATTTGTTATTTGAGTGTGCTTAA 181032
RESULT 33
US-09-902-540-8128
Sequence 8128, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10

NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 8128
LENGTH: 2455
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-8128
Query Match 5.9%; Score 35.2; DB 4; Length 2455;
Best Local Similarity 49.5%; Pred. No. 24;
Matches 91; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 149 GAGTGTGATGAGCTCTACAGAGTACAGGAGTACCCCAACAGAGAGGCA 208
DB 616 GAGAGGCTCTACAGAGTGTGAGGAGTGTGAGATTCGAGATTCGAGTGTGAGAGATC 675
QY 209 GAGAGATCATCAAGACCTCATCAAGAGATCATCAAGCTGGCCATCTTTATAGAG 268
DB 676 CCGAGCGCGCTCAAGAGAGATGTGAGAGAGAGAGAGAGAGTACTTACTTATAGAG 735
QY 269 AATCAGTTTATCAAGATGAGCTTACATTTGATGAGAGAAATTTAAGAGAGTTCATCAG 328
DB 736 ATGCAAGCCATTCAGAGAGAGCTGGGTGAGCGCGAGAGTTTCAAGAGAGATTCAGAG 795
QY 329 CTTG 332
DB 796 ATTG 799
RESULT 34
US-09-902-540-832/c
Sequence 832, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Gregory J.
APPLICANT: Slater, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 832
LENGTH: 8352
TYPE: DNA
ORGANISM: Myxococcus xanthus
US-09-902-540-832
Query Match 5.9%; Score 35.2; DB 4; Length 8352;
Best Local Similarity 49.5%; Pred. No. 3;
Matches 91; Conservative 0; Mismatches 93; Indels 0; Gaps 0;
QY 149 GAGTGTGATGAGCTCTACAGAGTACAGGAGTACCCCAACAGAGAGGCA 208
DB 4610 GAGAGCTCTACAGAGTGTGAGGAGTGTGAGATTCGAGATTCGAGTGTGAGAGATC 4551
QY 209 GAGAGATCATCAAGACCTCATCAAGAGATCATCAAGCTGGCCATCTTTATAGAGAT 268
DB 4550 CCGAGCGCGCTCAAGAGAGATGTGAGAGAGAGAGAGAGTACTTACTTATAGAGATC 4491
QY 269 AATCAGTTTATCAAGATGAGCTTACATTTGATGAGAGAAATTTAAGAGAGATTCATCAG 328
DB 4490 ATGCAAGCCATTCAGAGAGAGCTGGGTGAGCGCGAGAGTTCAGAGAGAGATTCAGAG 4431
QY 329 CTTG 332
DB 4430 ATTG 4427
RESULT 35
US-09-949-016-13146

; Sequence 13146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13146
; LENGTH: 58829
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1).. (58829)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13146

Query Match
Best Local Similarity 5.8%; Score 35.2; DB 4; Length 58829;
Matches 67; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 467 GTTAAATAGTCTTGTATCATTTTTCAGATTGGAATTTTGGCTGCTGTATATCCT 526
DB 57676 GTTGTGCAATCTCTAATGCAATTTTAACTTTAGATTGTGAATGACCTATGTGAAT 57735
DB 57736 TTTGTTTATTAACCCAGAAATTAATCAAGTTTAAATGTGTCAACAACTGTTCAT 57795

RESULT 36
US-09-949-016-194263
; Sequence 194263, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194263
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-194263

Query Match
Best Local Similarity 5.8%; Score 35; DB 4; Length 601;
Matches 95; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 93 TGGGTAATAGTGTCTCCAAATCCACGACCTTAATAGACGACAAAGTAGAGG 152
DB 231 TGTCTATATTATATCTAATCATGAGGCTCCGATATATTAAGCAAAATATGAGAGAG 290

QY 153 TGCTGATGAGCTCTACAGAGTGACCGAGGAGTACCCCAAAACAGAGGACAGAGA 212
DB 291 TGAAGCAAGATGATCAGACAGAAATTAATTAATCAACTTAAGACCCCATTTGCAATAC 350
QY 213 AGATCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGCGCATTTTATAGAAATATC 272
DB 351 AGAATGAAAATTAATCAAGAGATCAATTAAGAAAGCTTGGGCAACATTACAGACCAATC 410
QY 273 AGTTAATCAAGATG 287
DB 411 AATTTAACAGACATG 425

RESULT 37
US-09-949-016-17262/C
; Sequence 17262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17262
; LENGTH: 24428
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17262

Query Match
Best Local Similarity 5.8%; Score 35; DB 4; Length 24428;
Matches 95; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 93 TGGGTAATAGTGTCTCCAAATCCATCGCCACCTTAATAGACGACAAAGTAGAGG 152
DB 15975 TGTCTATATTATATTAATCAATCATGAGGCTCCGATATTAAGCAAAATATGAGAGAG 15916
QY 153 TGCTGATGAGCTCTACAGAGTGACCGAGGAGTACCCCAAAACAGAGGACAGAGA 212
DB 15915 TGAAGCAAGATGATCAGACAGAAATTAATTAATCAACTTAAGACCCCATTTGCAATAC 15856
QY 213 AGATCATCAAGAACCTCATCAAGACAGTCAATCAAGCTGCGCATTTTATAGAAATATC 272
DB 15855 AGAATGAAAATTAATCAAGAGATCAATTAAGAAAGCTTGGGCAACATTACAGACCAATC 15796
QY 273 AGTTAATCAAGATG 287
DB 15795 AATTTAACAGACATG 15781

RESULT 38
US-09-949-016-14033/C
; Sequence 14033, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 14033
; LENGTH: 784019
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(784019)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14033
```

```

Query Match          5.8%; Score 35; DB 4; Length 784019;
Best Local Similarity 61.5%; Pred. No. 86;
Matches 56; Conservative 0; Mismatches 35; Indels 0; Gaps 0;
```

```

QY 305 AATTAAAGAAAGTTGATGAGTTCATGACCGGTGATGTTCCATGAGTGTGAT 364
    |||||
DB 891 AATTAAAGAAAGTTGATGATATATAAGGGTTATTAACCTTGATGATCAATAGTCT 832
    |||||
QY 365 TATACCTTGACCGGAAATGTTATTCAGGC 395
    |||||
DB 831 TATACGTTTAAAGAAAGAAATGTAAACAAAGC 801
    |||||
```

```

RESULT 39
US-09-949-016-205410/C
; Sequence 205410, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 205410
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-205410
```

```

Query Match          5.8%; Score 34.8; DB 4; Length 601;
Best Local Similarity 58.8%; Pred. No. 1.4;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```

QY 221 AAGAACCTCATCAAGACATGATCGGCATCTTTATAGATAATCACTTTAAT 280
    |||||
DB 365 AAGGACACCATCAAGAAAGTGAAGAGAGCCGCCACAAGTTGAGAAATATCTGCAAT 306
    |||||
QY 281 CAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATT 322
    |||||
DB 305 CATGATAGCACTTTTGTATGTGTAACCTTGAGGAAGACATTT 264
    |||||
```

```

RESULT 40
US-09-949-016-205411/C
; Sequence 205411, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 205411
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-205411
```

```

Query Match          5.8%; Score 34.8; DB 4; Length 601;
Best Local Similarity 58.8%; Pred. No. 1.4;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```

QY 221 AAGAACCTCATCAAGACATGATCGGCATCTTTATAGATAATCACTTTAAT 280
    |||||
DB 534 AAGGACACCATCAAGAAAGTGAAGAGAGCCGCCACAAGTTGAGAAATATCTGCAAT 475
    |||||
QY 281 CAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATT 322
    |||||
DB 474 CATGATAGCACTTTTGTATGTGTAACCTTGAGGAAGACATTT 433
    |||||
```

```

RESULT 41
US-09-949-016-17549
; Sequence 17549, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17549
; LENGTH: 24345
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17549
```

```

Query Match          5.8%; Score 34.8; DB 4; Length 24345;
Best Local Similarity 58.8%; Pred. No. 13;
Matches 60; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
```

```

QY 221 AAGAACCTCATCAAGACATGATCGGCATCTTTATAGATAATCACTTTAAT 280
    |||||
DB 13643 AAGGACACCATCAAGAAAGTGAAGAGAGCCGCCACAAGTTGAGAAATATCTGCAAT 13702
    |||||
QY 281 CAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATT 322
    |||||
DB 13703 CATGATAGCACTTTTGTATGTGTAACCTTGAGGAAGACATTT 13744
    |||||
```

```

RESULT 42
US-09-908-988B-5
; Sequence 5, Application US/09908988B
; Patent No. 6740751
```

GENERAL INFORMATION:
APPLICANT: OLSON, ERIC
APPLICANT: SPENCER, JEFFREY A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR STABILIZING MICROTUBULES
TITLE OF INVENTION: IN STRIATED MUSCLE CELLS
FILE REFERENCE: MYO:028US
CURRENT APPLICATION NUMBER: US/09/908,988B
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/219,020
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 5
LENGTH: 1597
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: CDS
LOCATION: (299)..(1327)
US-09-908-988B-5

Query Match 5.8%; Score 34.6; DB 4; Length 1597;
Best Local Similarity 53.3%; Pred. No. 2.9;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 70 CGTTCAGGCAACAAAGAGTCTTGGTAAATGCTCCAAATCCATCCGCAACCTT 129
DB 784 CTTCACAGGACAGACAGCTGAGTAACTGCTCATCTGCTGCGGGGAGACGA 843
QY 130 AATAGACGACAAAGTGTAGGAGTGTGATGAGCTCTACAGAGTGCAGGAGTAC 189
DB 844 CCGAGTGCACAGCATCTCTCTCAGCTGAGAGACTGCTGAGAGTGCAGGAGATAG 903
QY 190 CCAAAACAAGAGAGG 206
DB 904 CCACGAGTGAAGAGG 920

RESULT 43
US-09-248-796A-80
Sequence 80, Application US/09248796A
Patent No. 6747137
GENERAL INFORMATION:
APPLICANT: Keith Weinstock et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
FILE REFERENCE: 107196.132
CURRENT APPLICATION NUMBER: US/09/248,796A
PRIOR FILING DATE: 1999-02-12
PRIOR APPLICATION NUMBER: US 60/074,725
PRIOR FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/096,409
PRIOR FILING DATE: 1998-08-13
NUMBER OF SEQ ID NOS: 28208
SEQ ID NO 80
LENGTH: 672
TYPE: DNA
ORGANISM: Candida albicans
US-09-248-796A-80

Query Match 5.7%; Score 34.4; DB 4; Length 672;
Best Local Similarity 52.0%; Pred. No. 2;
Matches 77; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 446 ACTGCCAAGTCACATGAGCGGTTAATATGCTTGGATCATTTTTCAGATGTGAATT 505
DB 112 ATTGATTAAGCATATGATTAAGTTGAACATTAATTTCAAGATTAATGGAAATGCAITCG 171
QY 506 TTGGCTGCTTGTATATATCTTTTGGATTTTAAACCCACCTTACAAAATATATGTGAT 565
DB 172 AAAAAGCGCGAAAGATATCACTTGGAGGATATATAACAAATTTTATGCAATTTGAAAT 231
QY 566 GGTATCAACAAATTTGATGAAGAGA 593

DB 232 ACTACGAACATCTGTTAATGAAAAGA 259

RESULT 44
US-09-854-133-448
Sequence 448, Application US/09854133
Patent No. 6759508
GENERAL INFORMATION:
APPLICANT: Lodes, Michael J.
APPLICANT: Mohamath, Raodoh
APPLICANT: Henderson, Robert A.
APPLICANT: Benson, Darin R.
APPLICANT: Secrist, Heather
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.475C10
CURRENT APPLICATION NUMBER: US/09/854,133
PRIOR FILING DATE: 2001-05-11
NUMBER OF SEQ ID NOS: 735
SOFTWARE: FastSeq for windows Version 3.0
SEQ ID NO 448
LENGTH: 677
TYPE: DNA
ORGANISM: Homo sapiens
US-09-854-133-448

Query Match 5.7%; Score 34.4; DB 4; Length 677;
Best Local Similarity 59.0%; Pred. No. 2;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 158 GATGAGCTTACAGAGTGCAGGAGTACACCCAAACAGAGAGGACAGAGATC 217
DB 193 GAGAGCTACGCGAGCGCTGGAAGATTGAGAGATCCAGAGAGGTGCGAGAAACGAG 252
QY 218 ATCAAGAACTCATCAAGCAGTCAATCAAGCTGGCCATTC 257
DB 253 ATGACACACTCTGTGAGACTGACATCGAGCAGGTGATGC 292

RESULT 45
US-09-949-016-430
Sequence 430, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIOR FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 430
LENGTH: 2349
TYPE: DNA
ORGANISM: Human
US-09-949-016-430

Query Match 5.7%; Score 34.4; DB 4; Length 2349;
Best Local Similarity 59.0%; Pred. No. 4.2;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 158 GATGAGCTTACAGAGTGCAGGAGTACACCCAAACAGAGAGGACAGAGATC 217
DB 556 GAGAGCTACGCGAGCGCTGGAAGATTGAGAGTCCAGAGAGTGCAGAAACGAGC 615

Qy 218 ATCAAGAACCTCATCAAGACAGTCAATCAAGCTGGCCATTG 257
| | | | | | | | | | | | | | | | | | | | | |
Db 616 ATGCACAACTCTCTGGAGACTGACATCGAGCAGGTGAGTC 655

Search completed: July 28, 2005, 19:03:54
Job time : 187 secs


```
C 81 36.4 6.1 8692 19 US-10-623-500-13 Sequence 13, Appl
C 82 36.4 6.1 98642 13 US-10-087-192-1105 Sequence 1105, Ap
C 83 36.2 6.0 759 18 US-10-425-115-137599 Sequence 137599,
C 84 36.2 6.0 2739 18 US-10-335-977-3199 Sequence 3199, Ap
C 85 36.2 6.0 3135 18 US-10-335-977-3200 Sequence 3200, Ap
C 86 36.2 6.0 28626 18 US-10-672-787-15 Sequence 15, Appl
C 87 36.2 6.0 33406 9 US-09-880-107-2380 Sequence 2380, Ap
C 88 36.2 6.0 40045 21 US-10-741-600-17795 Sequence 17795, A
C 89 36 6.0 828 18 US-10-424-559-109224 Sequence 109224,
C 90 35.8 6.0 272 9 US-09-864-761-18684 Sequence 18684, A
C 91 35.8 6.0 482 9 US-09-864-761-1935 Sequence 1935, Ap
C 92 35.8 6.0 49736 20 US-10-719-993-7050 Sequence 7050, Ap
C 93 35.8 6.0 153170 20 US-10-723-860-199 Sequence 199, Ap
C 94 35.8 6.0 185458 20 US-10-719-993-6864 Sequence 6864, Ap
C 95 35.8 6.0 253861 19 US-10-741-601-5611 Sequence 5611, Ap
C 96 35.8 6.0 261817 13 US-10-087-192-2002 Sequence 2002, Ap
C 97 35.8 6.0 301692 17 US-10-428-487-11 Sequence 11, Appl
C 98 35.8 6.0 310268 19 US-10-367-094-195 Sequence 195, Ap
C 99 35.6 5.9 1174 17 US-10-282-122A-23404 Sequence 23404, A
100 35.4 5.9 2455 8 US-08-813-323A-4 Sequence 4, Appl
```

ALIGNMENTS

RESULT 1

```
US-10-627-571-1
; Sequence 1, Application US/10627571
; Publication No. US20040082771A1
; GENERAL INFORMATION:
; APPLICANT: KASID, Usha N.
; APPLICANT: KUMAR, Deepak
; APPLICANT: GOKHALE, Prafulla
; APPLICANT: AHMAD, Imran
; TITLE OF INVENTION: ANTI-APOTOPIC GENE SCC-S2 AND DIAGNOSTIC AND THERAPEUTIC USES
; FILE REFERENCE: 223316
; CURRENT APPLICATION NUMBER: US/10/627,571
; CURRENT FILING DATE: 2003-07-25
; PRIOR APPLICATION NUMBER: US 60/264,062
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/US02/02212
; PRIOR FILING DATE: 2002-01-28
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1915
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: SCC-S2
US-10-627-571-1
```

Query Match 100.0%; Score 601; DB 18; Length 1915;

Best Local Similarity 100.0%; Pred. No. 3; 9e-169; Indels 0; Gaps 0;

Matches 601; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CTTAGCGTCCGGCGCGCGTCCGACCTCTCCGATGGCCACAGATGTCTTTAATCCAA 60
DB 100 CTTAGCGTCCGGCGCGCGTCCGACCTCTCCGATGGCCACAGATGTCTTTAATCCAA 159
QY 61 AAACCTGGCGCTTCAGGCAAAAGAAATCTTGGTAAATGGTCCAAATCCATCCG 120
DB 160 AAACCTGGCGCTTCAGGCAAAAGAAATCTTGGTAAATGGTCCAAATCCATCCG 219
QY 121 CACCACTTAAATAGACAGACAAGTAGTAGTGTGATGAGCTCTACAGATGACAG 180
DB 220 CACCACTTAAATAGACAGACAAGTAGTAGTGTGATGAGCTCTACAGATGACAG 279
QY 181 GGAATACACCAAAACAAAGAGAGGACAGAAATCATCAAGAACTCATCAAGACAGT 240
DB 280 GGAATACACCAAAACAAAGAGAGGACAGAAATCATCAAGAACTCATCAAGACAGT 339
```

```
QY 241 CATCAAGCTGGCATTCTTTAATAGGAATAATCAGTTAATCAAGATGAGCTAGCATGT 300
DB 340 CATCAAGCTGGCATTCTTTAATAGGAATAATCAGTTAATCAAGATGAGCTAGCATGT 399
QY 301 GGAGAAATTTAAGAAAGAAATTCATCAGCTTGTCTATGACCGTGTCACTTTCATCAGT 360
DB 400 GGAGAAATTTAAGAAAGAAATTCATCAGCTTGTCTATGACCGTGTCACTTTCATCAGT 459
QY 361 GGATTAATCCTTTGACCGGAATGTGTATCCAGGCTGTTAATGAATGACAGAGATGCT 420
DB 460 GGATTAATCCTTTGACCGGAATGTGTATCCAGGCTGTTAATGAATGACAGAGATGCT 519
QY 421 GGACCAATTCATTCAGCCGCCACCTCAGTCCAGACATGACAGGCTTAATTAATGCTCT 480
DB 520 GGACCAATTCATTCAGCCGCCACCTCAGTCCAGACATGACAGGCTTAATTAATGCTCT 579
QY 481 TGATCATTTTTCAGATTGGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTTAA 540
DB 580 TGATCATTTTTCAGATTGGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTTAA 639
QY 541 ACCCACTTACAAAACCTATGTGATGTATCAACAAAATGTTGATGAAGACATATG 600
DB 640 ACCCACTTACAAAACCTATGTGATGTATCAACAAAATGTTGATGAAGACATATG 699
QY 601 A 601
DB 700 A 700
```

RESULT 2

```
US-10-491-545A-48
; Sequence 48, Application US/10491545A
; Publication No. US20050130117A1
; GENERAL INFORMATION:
; APPLICANT: Chu, Peter
; APPLICANT: Li, Congfen
; APPLICANT: Liao, X. Charlene
; APPLICANT: Masuda, Esteban
; APPLICANT: Pardo, Jorge
; APPLICANT: Zhao, Haoran
; APPLICANT: Rigel Pharmaceuticals, Inc.
; TITLE OF INVENTION: Modulators of Lymphocyte Activation and Migration
; FILE REFERENCE: 021044-00030US
; CURRENT APPLICATION NUMBER: US/10/491,545A
; CURRENT FILING DATE: 2004-04-01
; PRIOR APPLICATION NUMBER: US 60/327,212
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: WO PCT/US02/31618
; PRIOR FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human TNF-induced protein G2-1 mRNA,
; OTHER INFORMATION: complete CDS
US-10-491-545A-48
```

Query Match 97.9%; Score 588.4; DB 22; Length 1892;

Best Local Similarity 99.7%; Pred. No. 2; 3e-165; Indels 1; Gaps 1;

Matches 600; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 CTTAGCGTCCGGCGCGCGTCCG-CGACTCTCCGATGGCCACAGATGTCTTTAATCCAA 59
DB 63 CTTAGCGTCCGGCGCGCGTCCGCACTCTCCGATGGCCACAGATGTCTTTAATCCAA 122
QY 60 AAACCTGGCGCTTCAGGCAAAAGAAATCTTGGTAAATGGTGTCCAAATCCATCCG 119
DB 123 AAACCTGGCGCTTCAGGCAAAAGAAATCTTGGTAAATGGTGTCCAAATCCATCCG 182
```

Oy	130	CCACCACCTTAATAGACGACACAGATGAGAGGCTGCTGATGAGCTTACAGAGTGACA	179
Db	183	CCACCACCTTAATAGACGACACAGATGAGAGGCTGCTGATGAGCTTACAGAGTGACA	242
Oy	180	GGGAGTACACCCAAAACAAAGAGAGGACAGAGAGATCATCAAGAACTTCATCAAGACG	239
Db	243	GGGAGTACACCCAAAACAAAGAGAGGACAGAGAGATCATCAAGAACTTCATCAAGACG	302
Oy	240	TCATCAAGCTGGCCATCTTTATAGAAATPATCAGTTTAATCAAGATGAGCTAGCATTTGA	299
Db	303	TCATCAAGCTGGCCATCTTTATAGAAATPATCAGTTTAATCAAGATGAGCTAGCATTTGA	362
Oy	300	TGGAGAAATTTAAGAGAAAGTTCAATCAGCTGTCTATGACCGGGTCAGTTTCATCAGG	358
Db	363	TGGAGAAATTTAAGAGAAAGTTCAATCAGCTGTCTATGACCGGGTCAGTTTCATCAGG	422
Oy	360	TGATATTATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAATGACAGAGATGC	419
Db	423	TGATATTATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAATGACAGAGATGC	482
Oy	420	TGCACCAAAATCATTACGCGCCACCTCAGCTGCCAAGTCACATGACGCGGTTAATATGCT	479
Db	483	TGCACCAAAATCATTACGCGCCACCTCAGCTGCCAAGTCACATGACGCGGTTAATATGCT	542
Oy	480	TTGATCATTTTTCAGATTGTGAAATTTTGGCTGCTGTATATCCTTTTGGAAATTTTA	539
Db	543	TTGATCATTTTTCAGATTGTGAAATTTTGGCTGCTGTATATCCTTTTGGAAATTTTA	602
Oy	540	AACCCCACTTACAAAACATCATGTGATGATCAACAAAATGTTGATGAAAGAACATAT	599
Db	603	AACCCCACTTACAAAACATCATGTGATGATCAACAAAATGTTGATGAAAGAACATAT	662
Oy	600	GA 601	
Db	663	GA 664	

RESULT 3

US-10-097-065-98

Sequence 98, Application US/10097065

Publication No. US20030055236A1

GENERAL INFORMATION:

APPLICANT: Moore, Paul A. et al.

TITLE OF INVENTION: 110 Human Secreted Proteins

FILE REFERENCE: P2021P1

CURRENT APPLICATION NUMBER: US/10/097, 065

CURRENT FILING DATE: 2002-03-14

PRIOR APPLICATION NUMBER: PCT/US98/27059

PRIOR FILING DATE: 1998-12-17

PRIOR APPLICATION NUMBER: 60/070, 923

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/068, 007

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/068, 057

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/068, 006

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/068, 369

PRIOR FILING DATE: 1997-12-19

PRIOR APPLICATION NUMBER: 60/068, 367

PRIOR FILING DATE: 1997-12-19

PRIOR APPLICATION NUMBER: 60/068, 368

PRIOR FILING DATE: 1997-12-19

PRIOR APPLICATION NUMBER: 60/068, 169

PRIOR FILING DATE: 1997-12-19

PRIOR APPLICATION NUMBER: 60/068, 053

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/068, 064

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/068, 054

PRIOR FILING DATE: 1997-12-18

PRIOR APPLICATION NUMBER: 60/068, 008

PRIOR FILING DATE: 1997-12-18

```

; PRIOR APPLICATION NUMBER: 60/066,365
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1926)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1928)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1934)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-097-065-98

Query Match      97.9%; Score 588.2; DB 14; Length 1943;
Best Local Similarity 99.3%; Pred. No. 2.7e-165;
Matches 590; Conservative 0; Mismatches 4; Indels 0; Gaps 0

QY      8 GTGCCGCGCGCGTGGCCGACCTCTCCGATGCGCCACAGATGCTTTAATCCAAAACCTG 67
DB      93 GTCCCGCGCGCGTGGCCGACCTCTCCGATGCGCCACAGATGCTTTAATCCAAAACCTG 152

QY      68 GCCGTTGACGACAAAAGAAAGATCTTGGGTAATAATGCTGTCMAATCCATCGCCACACC 127
DB      153 GCCGTTGACGACAAAAGAAAGATCTTGGGTAATAATGCTGTCMAATCCATCGCCACACC 212

QY      128 TTAAATAGACGACACAAGTAGTAGTGCTGGATGAGCTTACAGAGTAGACCGAGAGTAG 187
DB      213 TTAAATAGACGACACAAGTAGTAGTGCTGGATGAGCTTACAGAGTAGACCGAGAGTAG 272

QY      188 ACCCAAAACAAGAGAGAGGACGAGAGATCATCAAGAACTCATCAAGACATCATCAAG 247
DB      273 ACCCAAAACAAGAGAGAGGACGAGAGATCATCAAGAACTCATCAAGACATCATCAAG 332

QY      248 CTGGCCATTCTTTATAGGAATATCAAGTTTATCAAGATGAGCTAGATTGATGAGAA 307
DB      333 CTGGCCATTCTTTATAGGAATATCAAGTTTATCAAGATGAGCTAGATTGATGAGAA 392

QY      308 TTTAAGAAAGAAAGTTTCATCAGCTTGCTATGACCGTGGTCAAGTTTTCATCAGGTGGATTAT 367
DB      393 TTTAAGAAAGAAAGTTTCATCAGCTTGCTATGACCGTGGTCAAGTTTTCATCAGGTGGATTAT 452

QY      368 ACCCTTACCGGAAATGATGTTATCCAGCGCTGTAATGAATGACAGAGATGCTGCACAA 427
DB      453 ACCCTTACCGGAAATGATGTTATCCAGCGCTGTAATGAATGACAGAGATGCTGCACAA 512

QY      428 ATCATTCAGCGCCACTCATCTGCAAGTCACATGACGAGGTTAATATAGTCTTTGATCAT 487
DB      513 ATCATTCAGCGCCACTCATCTGCAAGTCACATGACGAGGTTAATATAGTCTTTGATCAT 572

QY      488 TTTTCAGATTGTGAATTTTGGCTGCTGTTATATCTTTTGGAAATTTTAAACCCAC 547
DB      573 TTTTCAGATTGTGAATTTTGGCTGCTGTTATATCTTTTGGAAATTTTAAACCCAC 632

QY      548 TTTACAAAACATATGATGATGATCAACAAAATGTTGATGTAAGAAACATATGA 601
DB      633 TTTACAAAACATATGATGATGATCAACAAAATGTTGATGTAAGAAACATATGA 686

RESULT 4
US-10-372-876-98
; Sequence 98, Application US/10372876
; Publication No. US20030204071A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Moore, Paul A. et al.
; TITLE OF INVENTION: 110 Human Secreted Proteins
; FILE REFERENCE: P2021P1
; CURRENT APPLICATION NUMBER: US/10/372,876
; CURRENT FILING DATE: 2003-02-26
; PRIOR APPLICATION NUMBER: 09/334,595
; PRIOR FILING DATE: 1999-06-17
; PRIOR APPLICATION NUMBER: PCT/US98/27059
; PRIOR FILING DATE: 1998-12-17
; PRIOR APPLICATION NUMBER: 60/070,923
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,007
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,057
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,006
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/068,369
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,367
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,368
; PRIOR FILING DATE: 1997-12-19
; PRIOR APPLICATION NUMBER: 60/068,169
; PRIOR FILING DATE: 1997-12-19
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 672
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 98
; LENGTH: 1943
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1926)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1928)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1934)
; OTHER INFORMATION: n equals a,t,g, or c
; US-10-372-876-98

```

Query Match 97.9%; Score 588.2; DB 17; Length 1943;

Best Local Similarity 99.3%; Pred. No. 2.7e-165; Matches 550; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 8 GTCCCGGCGCGCTGCGGCACTCTCCGATGCGCCACAGATGCTTTAATCCAAAACCTG 67
DB 93 GTCCCGGCGCGCTGCGGCACTCTCCGATGCGCCACAGATGCTTTAATCCAAAACCTG 152
QY 68 GCCGTTGAGGCAAAAAGAGATCTTGGGTAATGCTGCCAATTCATCGCCACACC 127
DB 153 GCCGTTGAGGCAAAAAGAGATCTTGGGTAATGCTGCCAATTCATCGCCACACC 212
QY 128 TTAATGAGAGCAACAAGTAGTGAGTGAGTGAAGCTCTACAGAGTGAACGAGGAGTAC 187
DB 213 TTAATGAGAGCAACAAGTAGTGAGTGAGTGAAGCTCTACAGAGTGAACGAGGAGTAC 272
QY 188 ACCCAAAAACAAGAGGAGGAGAGATCATCAAGACCTCATCAAGACAGTCATCAAG 247
DB 273 ACCCAAAAACAAGAGGAGGAGAGATCATCAAGACCTCATCAAGACAGTCATCAAG 332
QY 248 CTGGCATTCTTTATGAAATTAATCAATCAAGTAGAGTGAAGTGAAGTGAAGTGA 307
DB 333 CTGGCATTCTTTATGAAATTAATCAATCAAGTAGAGTGAAGTGAAGTGAAGTGA 392

```

```

QY 308 TTTAAGAAAGATTCATCAGCTTGCTATGACCGTGTCAGCTTTCATCAGGTGATTA 367
DB 393 TTTAAGAAAGATTCATCAGCTTGCTATGACCGTGTCAGCTTTCATCAGGTGATTA 452
QY 368 ACCTTGAACCGGAATGTGTTATCCAGGCTTTAAATGAATGACAGAGATGTCACCA 427
DB 453 ACCTTGAACCGGAATGTGTTATCCAGGCTTTAAATGAATGACAGAGATGTCACCA 512
QY 428 ATCATTACGCGCCACTCTGCTCCAAATGACATGACGAGGTTAAATGATCTTATCAT 487
DB 513 ATCATTACGCGCCACTCTGCTCCAAATGACATGACGAGGTTAAATGATCTTATCAT 572
QY 488 TTTTGAATTTGAAATTTTGGCTGCTGTAATTCCTTTGGAAATTTTAAACCCAC 547
DB 573 TTTTGAATTTGAAATTTTGGCTGCTGTAATTCCTTTGGAAATTTTAAACCCAC 632
QY 548 TTTAAGAAAGATTCATCAGCTTGCTATGACCGTGTCAGCTTTCATCAGGTGATTA 601
DB 633 TTTAAGAAAGATTCATCAGCTTGCTATGACCGTGTCAGCTTTCATCAGGTGATTA 686

```

RESULT 5

```

; US-10-755-889-215
; Sequence 215, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE REFERENCE: D0284 NP
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: US/10/755,889
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 215
; LENGTH: 1814
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-755-889-215

```

Query Match 93.9%; Score 564.4; DB 19; Length 1814;

Best Local Similarity 99.8%; Pred. No. 3.6e-158; Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 36 TGCCACAGATGCTTTAATTCAAAACCTGCGGTTCCAGGACAAAAGAGATCTTGG 95
DB 175 TGCCACAGATGCTTTAATTCAAAACCTGCGGTTCCAGGACAAAAGAGATCTTGG 234
QY 96 GTAAATGCTGTCAAATTCATGCGCACCTTAATGACGACAAAGTAGTGAGTGC 155
DB 235 GTAAATGCTGTCAAATTCATGCGCACCTTAATGACGACAAAGTAGTGAGTGC 294
QY 156 TGATGAGCTTACAGAGTGACCCAGGAGTACCCAAAACAAGAGGAGAGAGAGA 215
DB 295 TGATGAGCTTACAGAGTGACCCAGGAGTACCCAAAACAAGAGGAGAGAGAGA 354
QY 216 TCATCAAGAACTCTCAAGACAGTCATCAAGCTGCGCACTTTATGGAATTAATCAGT 275
DB 355 TCATCAAGAACTCTCAAGACAGTCATCAAGCTGCGCACTTTATGGAATTAATCAGT 414
QY 276 TTAATCAAGATGAGTAGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 335
DB 415 TTAATCAAGATGAGTAGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 474
QY 336 TGACGTTGCTCACTTTCATCAGGTGATTAATCCTTTGAACCGAATGTTATCCAGGC 395
DB 475 TGACGTTGCTCACTTTCATCAGGTGATTAATCCTTTGAACCGAATGTTATCCAGGC 534

```

Oy	396	GGTTAAATGAATGCAGAGATGCTGCACCAATCAATTAGGGCCACCTCAGCCAAGT	455
Db	535	TGTTAAATGAATGCAGAGATGCTGCACCAATCAATTAGGGCCACCTCAGCCAAGT	594
Oy	456	CACATGCA CGGGTTATAATATGCTCTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT	515
Db	595	CACATGCA CGGGTTATAATATGCTCTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT	654
Oy	516	TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACCAAAACTATGTGATGATCAACA	575
Db	655	TGTATTAATCCTTTTGGGAATTTTAAACCCCACTTACCAAAACTATGTGATGATCAACA	714
Oy	576	AAATGTGGATGAAGAGAACTATGA	601
Db	715	AAATGTGGATGAAGAGAACTATGA	740

RESULT 6

```

US-10-755-889-217
/ Sequence 217, Application US/10755889
/ Publication No. US20040171823A1
/ GENERAL INFORMATION:
/ APPLICANT: Bristol-Myers Squibb Company
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
/ FILE REFERENCE: D0284 NP
/ CURRENT APPLICATION NUMBER: US/10/755,889
/ CURRENT FILING DATE: 2004-01-13
/ PRIOR APPLICATION NUMBER: U.S. 60/440,068
/ PRIOR FILING DATE: 2003-01-14
/ PRIOR APPLICATION NUMBER: U.S. 60/469,757
/ PRIOR FILING DATE: 2003-05-12
/ NUMBER OF SEQ ID NOS: 823
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 217
/ LENGTH: 2003
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-755-889-217

```

Query Match	93.9%	Score 564.4	DB 19	Length 2003
Best Local Similarity	99.8%	Pred. No. 3.9e-158		
Matches 565	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	36	TGCGCACAGATGTCCTTTAAATCCAAAACTGGCGGTTCAAGGCACAAAAGAAGATCTTGG	95	
Db	205	TGGCCACAGATGTCCTTTAAATCCAAAACTGGCGGTTCAAGGCACAAAAGAAGATCTTGG	264	
QY	96	GTAATAATGCTGCCAATTCATTCGCCACACCTTAATGACGACACAAATGTGAGGTGC	155	
Db	265	GTAATAATGCTGCCAATTCATTCGCCACACCTTAATGACGACACAAATGTGAGGTGC	324	
QY	156	TGATGAGCTCTACAGAGTGACCAAGGAGTACACCCAAAACAAGAGGACAGAAATG	215	
Db	325	TGATGAGCTCTACAGAGTGACCAAGGAGTACACCCAAAACAAGAGGACAGAAATG	384	
QY	216	TCATCAAGAACTCATCAAGACAGTCATCAAGCTGGCCATCTTTAATGAAATATCAGT	275	
Db	385	TCATCAAGAACTCATCAAGACAGTCATCAAGCTGGCCATCTTTAATGAAATATCAGT	444	
QY	276	TTAATCAAGATGAGCTAGCATTTGATGGGAATTTAAGAAAGAAAGTTATCAGCTGCTA	335	
Db	445	TTAATCAAGATGAGCTAGCATTTGATGGGAATTTAAGAAAGAAAGTTATCAGCTGCTA	504	
QY	336	TGACCGTGTCAGTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTATTCAAGC	395	
Db	505	TGACCGTGTCAGTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTATTCAAGC	564	
QY	396	TGTTAAATGAATGACAGAGATGCTGCACCAATTCATTCAAGGCACTCATCTGCAAGT	455	
Db	565	TGTTAAATGAATGACAGAGATGCTGCACCAATTCATTCAAGGCACTCATCTGCAAGT	624	
QY	456	CACATGACGGGTTAATATGTCCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCTT	515	

Accession	Sequence	Position
Db	CGATGGACCGGTTAATTAATGCTTGAATCTTTTCAGATTGGAAATTTGGCTGCT	68
Oy	TGTAATATCCTTTTGGGAATTTAAACCCACTTACAAAATAATGATGATCAACA	575
Db	TGTAATATCCTTTTGGGAATTTTAAACCCCACTTACAAAATAATGATGATCAACA	744
Oy	AAATGTGGATGAAGGAACATATGA	601
Db	AAATGTTGATGAAGGAACATATGA	770

RESULT 7

```

US-10-087-192-497
? Sequence 497, Application US/10087192
? Publication No. US20020182586A1
? GENERAL INFORMATION:
? APPLICANT: Morris, David W.
? APPLICANT: Engelhard, Eric K.
? TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
? TITLE OF INVENTION: CANCER
? FILE REFERENCE: 529452000122
? CURRENT APPLICATION NUMBER: US/10/087,192
? CURRENT FILING DATE: 2002-03-01
? PRIOR APPLICATION NUMBER: US 09/747,377
? PRIOR FILING DATE: 2000-12-22
? PRIOR APPLICATION NUMBER: US 09/798,586
? PRIOR FILING DATE: 2001-03-02
? NUMBER OF SEQ ID NOS: 2059
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 497
? LENGTH: 2034
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-087-192-497

```

Query Match	93.9%	Score 564.4	DB 13	Length 2034
Best Local Similarity	99.8%	Pred. No. 3.9e-158		
Matches 565	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	36	TGACCACAGATGTCCTTTAATTC	CAAAAAACCTGGCCGTTCAAGGCACAAAAGAGATCTTGG	95
Db	249	TGGCCACAGATGTCCTTTAATTC	CAAAAAACCTGGCCGTTCAAGGCACAAAAGAGATCTTGG	308
Qy	96	GTAATAATGCTGTCCAAATCCATG	CCACCTTAAATGAGACACAGATGATGAGTGC	155
Db	309	GTAATAATGCTGTCCAAATCCATG	CCACCTTAAATGAGACACAGATGATGAGTGC	368
Qy	156	TGATGAGCTCTACAGAGTGAC	CCAGGAGTACACCCAAAACAAGAGGCGACAGAA	215
Db	369	TGATGAGCTCTACAGAGTGAC	CCAGGAGTACACCCAAAACAAGAGGCGACAGAA	428
Qy	216	TCATCAGAACCCTCATTAAGAC	AGTCAATCAAGCTGGCCATCTTTAATAGAAATATCAGT	275
Db	429	TCATCAGAACCCTCATTAAGAC	AGTCAATCAAGCTGGCCATCTTTAATAGAAATATCAGT	488
Qy	276	TTAATCAAGATGAGCTACGATT	GAATGAGAAATTTAAGAAAGAGTTCATCAGCTTGCTA	335
Db	489	TTAATCAAGATGAGCTACGATT	GAATGAGAAATTTAAGAAAGAGTTCATCAGCTTGCTA	548
Qy	336	TGACCGTGTCAGTTTCATCAG	GTGATTAACCTTTGACCCGAAATGTGTTATCCAGGC	395
Db	549	TGACCGTGTCAGTTTCATCAG	GTGATTAACCTTTGACCCGAAATGTGTTATCCAGGC	608
Qy	396	TGTTAAATGAAATGACAGAGA	TGCTGTCACCAATCATTCACGCGCCACTCTGCCAAGT	455
Db	609	TGTTAAATGAAATGACAGAGA	TGCTGTCACCAATCATTCACGCGCCACTCTGCCAAGT	668
Qy	456	CACATGACCGGGTTAATPAAT	GTCTTTGATCATTCTTTTTCAGATTGGAATTTTGGCTGCT	515
Db	669	CACATGACCGGGTTAATPAAT	GTCTTTGATCATTCTTTTTCAGATTGGAATTTTGGCTGCT	728
Qy	516	TGTATATCTCTTTGGGAATTT	TAAACCCCACTTAACAAAACATAATGATGATATCAACA	575


```

US-10-029-386-22914
; Sequence 22914, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine, vers. 1.1
; SEQ ID NO 22914
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC035144.2 SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: EST HUMAN HIT: AL549492.1, EVALU0 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: P31390, EVALU0 2.60e-01
; OTHER INFORMATION: NT HIT: g17657123, EVALU0 0.00e+00
US-10-029-386-22914

```

Query Match 90.2%; Score 542.4; DB 16; Length 544;
Best Local Similarity 99.8%; Pred. No. 7.1e-152;
Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 36 TGCCACAGATGCTTATTCCTTAAATTCCTGCGCTTCAAGGACCAAAAGAGATCTTGG 95
DB 1 TGCGCCAGATGCTTATTCCTTAAATTCCTGCGCTTCAAGGACCAAAAGAGATCTTGG 60
QY 96 GTTAAATGCTGCTCAATTCATGCGCCACCTTATATGACGACACAAAGTATGAGGTGC 155
DB 61 GTTAAATGCTGCTCAATTCATGCGCCACCTTATATGACGACACAAAGTATGAGGTGC 120
QY 156 TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCAAACAAAGAGAGGACAGAGA 215
DB 121 TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCAAACAAAGAGAGGACAGAGA 180
QY 216 TCATCAAGAACTCATCAAGACAGTCAAGCTGCGCATCTTTATATGAGATATCATG 275
DB 181 TCATCAAGAACTCATCAAGACAGTCAAGCTGCGCATCTTTATATGAGATATCATG 240
QY 276 TTAATCAAGATGAGCTTATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 335
DB 241 TTAATCAAGATGAGCTTATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 300
QY 336 TGAACCTGCTGATGCTTTCATCAGGTGATTAATCTTTGACCGGATGTTATCCAGGC 395
DB 301 TGAACCTGCTGATGCTTTCATCAGGTGATTAATCTTTGACCGGATGTTATCCAGGC 360
QY 396 TGTAAATGATGATGAGAGATGCTGACCAAAATCATTCAGCGGACCTCATGCGCAAGT 455
DB 361 TGTAAATGATGATGAGAGATGCTGACCAAAATCATTCAGCGGACCTCATGCGCAAGT 420
QY 456 CACATGAGCGGTTAATATATGCTTGTGATCATTTTTCAGATTTTGGTGGCT 515
DB 421 CACATGAGCGGTTAATATATGCTTGTGATCATTTTTCAGATTTTGGTGGCT 480
QY 516 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACTATGATGATGATGATG 575
DB 481 TGTATATCTTTTGGGAATTTTAAACCCACTTACAAAACTATGATGATGATGATG 540
QY 576 AAAT 579
DB 541 AAAT 544

```

```

RESULT 11
US-10-087-192-493
; Sequence 493, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morrie, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 493
; LENGTH: 62231
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(62231)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-493

```

Query Match 74.2%; Score 446; DB 13; Length 62231;
Best Local Similarity 86.1%; Pred. No. 1.1e-121;
Matches 494; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

```

QY 28 TCCTCCAGTGGCCACAGATGCTTATTCCTTAAATTCCTGCGCTTCAAGGACCAAAAGAA 87
DB 50644 TTCTGCAATGCTGCTACAGATGCTTATTCCTTAAATTCCTGCGCTTCAAGGACCAAAAGAA 50703
QY 88 GATCTTGGGTTAAATGCTGCTCAATTCATGCGCCACCTTATATGACGACCAAGTAG 147
DB 50704 GATCTTGGGTTAAATGCTGCTCAATTCATGCGCCACCTTATATGACGACCAAGTAG 50763
QY 148 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 207
DB 50764 TGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 50823
QY 208 AGAAGATCATCAAGAACTCATCAAGACAGTCAAGCTGCGCATCTTTATATGAGAA 267
DB 50824 AGAAGATCATCAAGAACTCATCAAGACAGTCAAGCTGCGCATCTTTATATGAGAA 50883
QY 268 TATCAGTTTAAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 327
DB 50884 CATCAGTTTAAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 50943
QY 328 GCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 387
DB 50944 GCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 51003
QY 388 ATCCAGGCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 447
DB 51004 ATCCAGGCTGTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 51063
QY 448 TGCAAGTCAATGAGCGGTTAATATATGCTTGTGATCATTTTTCAGATTTTGGTGGCT 507
DB 51064 TGCAAGTCAATGAGCGGTTAATATATGCTTGTGATCATTTTTCAGATTTTGGTGGCT 51123
QY 508 GGTGCTGCTGATATATCTTTTGGGAATTTTAAACCCACTTACAAAACTATGATGATG 567
DB 51124 GGTGCTGCTGATATATCTTTTGGGAATTTTAAACCCACTTACAAAACTATGATGATG 51183
QY 568 TATCACAAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601
DB 51184 TATCACAAAATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 51217

```


RESULT 12
 US-10-087-192-494
 ; Sequence 494, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 494
 ; LENGTH: 2087
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 ; US-10-087-192-494

Query Match 73.9%; Score 444.4; DB 13; Length 2087;
 Best Local Similarity 86.6%; Pred. No. 4.2e-122; Indels 0; Gaps 0;
 Matches 490; Conservative 0; Mismatches 76;

QY 36 TGGCCACAGATGCTTTAAATTCAAAACCTGGCCGTTCCAGGACAAAAGATCTTGG 95
 DB 508 TGGCTACAGATGCTTCAATTCAAAACCTGGCCGTTCCAGGACAAAAGATCTTGG 567
 QY 96 GTAAATGGTGCTCAATTCATGCCACCACTTAAATAGACACAGATGAGGTGC 155
 DB 568 GCAAAATGGATCCAATTCATGCCACCACTGATGACGACACAGACGAGGTGC 627
 QY 156 TGGATAGCTTACAGATGACAGGAGTACCCAAAACAAGAGGAGGAGAGAA 215
 DB 628 TGAATAGCTTACAGGAGTACCAAGAGTACCCAAAACAAGAGGAGGAGAGG 687
 QY 216 TCATCAAGAACTCATCAAGACAGTCAAGCTGGCCATTTTATAGGAATATCAGT 275
 DB 688 TCATCAAGAACTCATCAAGAGGTGATCAAGCTGGCCCTCCACAGGAATATCAGT 747
 QY 276 TTAATCAAGATGCTAGATGATGAGAAATTTAAGAAAGTTTATCAGCTTGTCTA 335
 DB 748 TCAATCAAGACGAGCTGGCGCTCATGAGAAATTTCAAGAAAGTGCACAGCTTGCCA 807
 QY 336 TGAACGTCGTCAGTTTCCATCAGGTGATTAATCTTTGACCGGAATGTTATCCAGGC 395
 DB 808 TGAACGTCGTCAGCTTCCACAGGTGATTAATCTTTGACCGGAATGTCAGAGC 867
 QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTCAGCCGACCTCATGCCAAGT 455
 DB 868 TGTGAACAGATGCTGCAGAGCTCTTACAGAGATCATTCAGCCGACCTTACGCCAAGT 927
 QY 456 CACATGAGCGGTTAATTAATGTCCTTGATCATTTTTCAGATGTAATTTTGGCTGCT 515
 DB 928 CTCACGAGCGGTTAATTAATGTCCTTGATCATTTTTCAGATGTAATTTTGGCTGCT 987
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCCATTAAGAAAATGATGATGATTAACA 575
 DB 988 TGTACATCTCTTTGGAAGTTTAAACCTCATTTACGAAACTTTGGAGCGGATCAACA 1047
 QY 576 AAATGTTGATGAAGAAACATATGA 601
 DB 1048 AAATGTTGATGAAGAAACATATGA 1073

RESULT 13
 US-09-796-692-3207

; Sequence 3207, Application US/09796692
 ; Publication No. US20020198362A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gaiger, Alexander
 ; APPLICANT: Algate, Paul A.
 ; APPLICANT: Mannion, Jane
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
 ; FILE REFERENCE: 2077 001200
 ; CURRENT APPLICATION NUMBER: US/09/796,692
 ; CURRENT FILING DATE: 2001-03-01
 ; PRIOR APPLICATION NUMBER: 60/186,126
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: 60/190,479
 ; PRIOR FILING DATE: 2000-03-17
 ; PRIOR APPLICATION NUMBER: 60/200,545
 ; PRIOR FILING DATE: 2000-04-27
 ; PRIOR APPLICATION NUMBER: 60/200,303
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: 60/200,779
 ; PRIOR FILING DATE: 2000-04-28
 ; PRIOR APPLICATION NUMBER: 60/200,999
 ; PRIOR FILING DATE: 2000-05-01
 ; PRIOR APPLICATION NUMBER: 60/202,084
 ; PRIOR FILING DATE: 2000-05-04
 ; PRIOR APPLICATION NUMBER: 60/206,201
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: 60/218,950
 ; PRIOR FILING DATE: 2000-07-14
 ; PRIOR APPLICATION NUMBER: 60/222,903
 ; PRIOR FILING DATE: 2000-08-03
 ; PRIOR APPLICATION NUMBER: 60/223,416
 ; PRIOR FILING DATE: 2000-08-04
 ; PRIOR APPLICATION NUMBER: 60/223,378
 ; PRIOR FILING DATE: 2000-08-07
 ; NUMBER OF SEQ ID NOS: 9597
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 3207
 ; LENGTH: 287
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-796-692-3207

Query Match 47.5%; Score 285.4; DB 9; Length 287;
 Best Local Similarity 99.7%; Pred. No. 6.7e-75;
 Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 252 CCATCTTTATAGAAATATCAATTCAGTTTATCAAGATGAGCTAGCATTTGAGAAATTTA 311
 DB 1 CCATCTTTATAGAAATATCAATTCAGTTTATCAAGATGAGCTAGCATTTGAGAAATTTA 60
 QY 312 AGAAGAAATTCATCAGCTTGTCTATGACCGTGTCACTTTCCATCAGGTGATTAATCCT 371
 DB 61 AGAAGAAATTCATCAGCTTGTCTATGACCGTGTCACTTTCCATCAGGTGATTAATCCT 120
 QY 372 TTGACCGAATGTTATTCAGGCTGTTAAATGAATGAGAGATGCTGCACCAATCA 431
 DB 121 TTGACCGAATGTTATTCAGGCTGTTAAATGAATGAGAGATGCTGCACCAATCA 180
 QY 432 TTCAGCGCACCTCATCTGCACATGACATGACGAGGTTAATATATGCTTTGATCAATTTT 491
 DB 181 TTCAGCGCACCTCATCTGCACATGACATGACGAGGTTAATATATGCTTTGATCAATTTT 240
 QY 492 CAGATTGTAATTTTGGCTGCTGCTTGTATATCTTTTGGGAATTT 538
 DB 241 CAGATTGTAATTTTGGCTGCTGCTTGTATATCTTTTGGGAATTT 287

RESULT 14
 US-10-040-862-3207
 ; Sequence 3207, Application US/10040862
 ; Publication No. US20030078396A1
 ; GENERAL INFORMATION:

APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Reiter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/040,862
CURRENT FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: US 60/223,416
PRIOR FILING DATE: 2000-08-04
PRIOR APPLICATION NUMBER: US 60/223,378
PRIOR FILING DATE: 2000-08-07
PRIOR APPLICATION NUMBER: US 09/796,692
PRIOR FILING DATE: 2001-03-01
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3207
LENGTH: 287
TYPE: DNA
ORGANISM: Homo sapiens
US-10-040-862-3207

Query Match 47.5%; Score 285.4; DB 14; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 252 CCATTCTTTATAGAAATTAATCAAGTTAATCAAGTACAGTTCGAGGATTAACCT 311
DB 1 CCATTCTTTATAGAAATTAATCAAGTTAATCAAGTACAGTTCGAGGATTAACCT 60
QY 312 AGAAGAAAGTTATCATGCTGCTATGACCGTGTGCTGTTCCATCAGGTGATTAACCT 371
DB 61 AGAAGAAAGTTATCATGCTGCTATGACCGTGTGCTGTTCCATCAGGTGATTAACCT 120
QY 372 TTGACCGGAATGTATTACAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB 121 TTGACCGGAATGTATTACAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATCA 180
QY 432 TTGACCGCACTCATGCTGCAAGTCAATGACAGGCTTAATATGCTTTGATCAATTTT 491
DB 181 TTGACCGCACTCATGCTGCAAGTCAATGACAGGCTTAATATGCTTTGATCAATTTT 240
QY 492 CAGATTGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTT 538
DB 241 CAGATTGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTT 287

RESULT 15
US-10-057-475B-3207
; Sequence 3207, Application US/10057475B
; Publication No. US20040002068A1

GENERAL INFORMATION:
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Clapper, Jonathan David
APPLICANT: Wang, Ai-jun
APPLICANT: Ordenez, Nadia
APPLICANT: Carter, Lauren
APPLICANT: McNeill, Patricia Dianne
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-014402US
CURRENT APPLICATION NUMBER: US/10/057,475B
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3207
LENGTH: 287
TYPE: DNA
ORGANISM: Homo sapiens
US-10-057-475B-3207

Query Match 47.5%; Score 285.4; DB 17; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 252 CCATTCTTTATAGAAATTAATCAAGTTAATCAAGTACAGTTCGAGGATTAACCT 311
DB 1 CCATTCTTTATAGAAATTAATCAAGTTAATCAAGTACAGTTCGAGGATTAACCT 60
QY 312 AGAAGAAAGTTATCATGCTGCTATGACCGTGTGCTGTTCCATCAGGTGATTAACCT 371
DB 61 AGAAGAAAGTTATCATGCTGCTATGACCGTGTGCTGTTCCATCAGGTGATTAACCT 120
QY 372 TTGACCGGAATGTATTACAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB 121 TTGACCGGAATGTATTACAGGCTGTTAAATGAATGACAGAGATGCTGCACCAATCA 180
QY 432 TTGACCGCACTCATGCTGCAAGTCAATGACAGGCTTAATATGCTTTGATCAATTTT 491
DB 181 TTGACCGCACTCATGCTGCAAGTCAATGACAGGCTTAATATGCTTTGATCAATTTT 240
QY 492 CAGATTGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTT 538
DB 241 CAGATTGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTT 287

RESULT 16
US-10-154-884B-3207
; Sequence 3207, Application US/10154884B
; Publication No. US20040005361A1

```
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannon, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-154-884B-3207

Query Match          47.5%; Score 285.4; DB 17; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 252 CCATCTTTATAGAAATATCAAGTTAAATCAAGATGAGATGAGATGAGAAATTTA 311
DB 1 CCATCTTTATAGAAATATCAAGTTAAATCAAGATGAGATGAGATGAGAAATTTA 60
QY 312 AGAAGAAAGTTATATCAGCTTGCTATGACCGTGTGCTAGTTCCATCAGGTGATTAACCT 371
DB 61 AGAAGAAAGTTATATCAGCTTGCTATGACCGTGTGCTAGTTCCATCAGGTGATTAACCT 120
QY 372 TTGACCGGAATGTATATCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB 121 TTGACCGGAATGTATATCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 180
QY 432 TTCAGGCCCACTCACTGCGCAAGTCAATGACAGCGGTTAATATATGCTTTGATCATTTT 491
DB 181 TTCAGGCCCACTCACTGCGCAAGTCAATGACAGCGGTTAATATGCTTTGATCATTTT 240
QY 492 CAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 538
DB 241 CAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 287

RESULT 17
US-10-764-324-3207
/ Sequence 3207, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannon, Jane
```

```
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ CURRENT FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 3207
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-764-324-3207

Query Match          47.5%; Score 285.4; DB 19; Length 287;
Best Local Similarity 99.7%; Pred. No. 6.7e-75;
Matches 286; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 252 CCATCTTTATAGAAATATCAAGTTAAATCAAGATGAGATGAGATGAGAAATTTA 311
DB 1 CCATCTTTATAGAAATATCAAGTTAAATCAAGATGAGATGAGATGAGAAATTTA 60
QY 312 AGAAGAAAGTTATATCAGCTTGCTATGACCGTGTGCTAGTTCCATCAGGTGATTAACCT 371
DB 61 AGAAGAAAGTTATATCAGCTTGCTATGACCGTGTGCTAGTTCCATCAGGTGATTAACCT 120
QY 372 TTGACCGGAATGTATATCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 431
DB 121 TTGACCGGAATGTATATCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATCA 180
QY 432 TTCAGGCCCACTCACTGCGCAAGTCAATGACAGCGGTTAATATATGCTTTGATCATTTT 491
DB 181 TTCAGGCCCACTCACTGCGCAAGTCAATGACAGCGGTTAATATGCTTTGATCATTTT 240
QY 492 CAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 538
DB 241 CAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 287

RESULT 18
US-09-796-692-2801/C
/ Sequence 2801, Application US/09796692
/ Publication No. US20020198362A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannon, Jane
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
/ TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
/ FILE REFERENCE: 2077.001200
/ CURRENT APPLICATION NUMBER: US/09/796,692
```

```

; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-796-692-2801
```

```

Query Match      47.2%; Score 283.8; DB 9; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 252 CCATCTTTATAGGAATATCAAGTATGAGTATGATGAGAAATTTA 311
DB 287 CCATCTTTATAGGAATATCAAGTATGAGTATGATGAGAAATTTA 228
QY 312 AGAAGAAAGTTATCATGCTGTATGACCGTGTGCTGATTCATCAGGTGATTAACCT 371
DB 227 AGAAGAAAGTTATCATGCTGTATGACCGTGTGCTGATTCATCAGGTGATTAACCT 168
QY 372 TTGACCGGAATGTATATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTATATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATCA 108
QY 432 TTGACCGCACTCACTGCAAGTCAATGACAGCGGTTAATATGCTTTGATCAATTTT 491
DB 107 TTGACCGCACTCACTGCAAGTCAATGACAGCGGTTAATATGCTTTGATCAATTTT 48
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 1
```

```

RESULT 19
US-10-040-862-2801/C
; Sequence 2801, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
```

```

; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2801
; LENGTH: 287
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-040-862-2801
```

```

Query Match      47.2%; Score 283.8; DB 14; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 252 CCATCTTTATAGGAATATCAAGTATGAGTATGATGAGAAATTTA 311
DB 287 CCATCTTTATAGGAATATCAAGTATGAGTATGATGAGAAATTTA 228
QY 312 AGAAGAAAGTTATCATGCTGTATGACCGTGTGCTGATTCATCAGGTGATTAACCT 371
DB 227 AGAAGAAAGTTATCATGCTGTATGACCGTGTGCTGATTCATCAGGTGATTAACCT 168
QY 372 TTGACCGGAATGTATATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTATATCCAGGCTGTTAATGAATGACAGAGATGCTGCACCAATCA 108
QY 432 TTGACCGCACTCACTGCAAGTCAATGACAGCGGTTAATATGCTTTGATCAATTTT 491
DB 107 TTGACCGCACTCACTGCAAGTCAATGACAGCGGTTAATATGCTTTGATCAATTTT 48
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATCCTTTTGGGAATTTT 1
```

```

RESULT 20
US-10-057-475B-2801/C
; Sequence 2801, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Galger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aljun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
```

```
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-014402US
/ CURRENT APPLICATION NUMBER: US/10/057,475B
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 10979
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2801
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-057-475B-2801
```

```
Query Match          47.2%; Score 283.8; DB 17; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 252 CCATTCTTATAGGAATATACGTTTAATCAAGATGAGTATGATGAGAAATTTA 311
DB 287 CCAATCTTTATAGGATATACGTTTAATCAAGATGAGTATGATGAGAAATTTA 228
QY 312 AGAAGAAAGTTATACGCTTGTATGACCGTGTGCTGATGCTTCCATCAGTGGATTAACCT 371
DB 227 AGAAGAAAGTTATACGCTTGTATGACCGTGTGCTGATGCTTCCATCAGTGGATTAACCT 168
QY 372 TTGACCGGAATGTATTATCCAGGCTGTATTAATGAATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTATTATCCAGGCTGTATTAATGAATGACAGAGATGCTGCACCAATCA 108
QY 432 TTCAGCGCACCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTGATCATTTT 491
DB 107 TTCAGCGCACCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTGATCATTTT 48
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 1
```

```
RESULT 21
US-10-154-884B-2801/C
/ Sequence 2801, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B
```

```
/ CURRENT FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2801
/ LENGTH: 287
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-154-884B-2801
```

```
Query Match          47.2%; Score 283.8; DB 17; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 252 CCATTCTTATAGGAATATACGTTTAATCAAGATGAGTATGATGAGAAATTTA 311
DB 287 CCAATCTTTATAGGATATACGTTTAATCAAGATGAGTATGATGAGAAATTTA 228
QY 312 AGAAGAAAGTTATACGCTTGTATGACCGTGTGCTGATGCTTCCATCAGTGGATTAACCT 371
DB 227 AGAAGAAAGTTATACGCTTGTATGACCGTGTGCTGATGCTTCCATCAGTGGATTAACCT 168
QY 372 TTGACCGGAATGTATTATCCAGGCTGTATTAATGAATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTATTATCCAGGCTGTATTAATGAATGACAGAGATGCTGCACCAATCA 108
QY 432 TTCAGCGCACCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTGATCATTTT 491
DB 107 TTCAGCGCACCTCACTGCAAGTCAATGACAGGCTTAATATGCTTTGATCATTTT 48
QY 492 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATATCCTTTGGGAATTT 1
```

```
RESULT 22
US-10-764-324-2801/C
/ Sequence 2801, Application US/10764324
/ Publication No. US20040175739A1
/ GENERAL INFORMATION:
/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013520US
/ CURRENT APPLICATION NUMBER: US/10/764,324
/ CURRENT FILING DATE: 2004-01-23
/ PRIOR APPLICATION NUMBER: US/10/040,862
/ PRIOR FILING DATE: 2001-11-06
/ PRIOR APPLICATION NUMBER: US 60/186,126
```

```
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: US 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: US 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: US 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,779
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: US 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: US 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: US 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: US 60/218,950
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 10467
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 2801
;; LENGTH: 287
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-10-764-324-2801

Query Match      47.2%; Score 283.8; DB 19; Length 287;
Best Local Similarity 99.3%; Pred. No. 2e-74;
Matches 285; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 252 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGTTCAGTTCATGAGAAATTTA 311
DB 287 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGTTCAGTTCATGAGAAATTTA 228

QY 312 AGAAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTTCCATCAGTGGATTATCT 371
DB 227 AGAAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTTCCATCAGTGGATTATCT 168

QY 372 TTGACCGGAATGTGTATTCAGGCTGTAAATGATGACAGAGATGCTGCACCAATCA 431
DB 167 TTGACCGGAATGTGTATTCAGGCTGTAAATGATGACAGAGATGCTGCACCAATCA 108

QY 432 TTGACCGGAATGTGTATTCAGGCTGTAAATGATGACAGAGATGCTGCACCAATCA 491
DB 107 TTGACCGGAATGTGTATTCAGGCTGTAAATGATGACAGAGATGCTGCACCAATCA 48

QY 492 CAGATTGTGAATTTTGGCTGCTGTATATCTTTTGGAAATTTT 538
DB 47 CAGATTGTGAATTTTGGCTGCTGTATATCTTTTGGAAATTTT 1

RESULT 23
US-09-796-692-5724
;; Sequence 5724, Application US/09796692
;; Publication No. US20020198362A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY
;; TITLE OF INVENTION: HEMATOLOGICAL MALIGNANCIES
;; FILE REFERENCE: 2077,001200
;; CURRENT APPLICATION NUMBER: US/09/796,692
;; PRIOR FILING DATE: 2001-03-01
;; PRIOR APPLICATION NUMBER: 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: 60/190,479
;; PRIOR FILING DATE: 2000-03-17
;; PRIOR APPLICATION NUMBER: 60/200,545
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 60/200,303
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,779
```

```
;; PRIOR FILING DATE: 2000-04-28
;; PRIOR APPLICATION NUMBER: 60/200,999
;; PRIOR FILING DATE: 2000-05-01
;; PRIOR APPLICATION NUMBER: 60/202,084
;; PRIOR FILING DATE: 2000-05-04
;; PRIOR APPLICATION NUMBER: 60/206,201
;; PRIOR FILING DATE: 2000-05-22
;; PRIOR APPLICATION NUMBER: 60/218,950
;; PRIOR FILING DATE: 2000-07-14
;; PRIOR APPLICATION NUMBER: 60/222,903
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: 60/223,416
;; PRIOR FILING DATE: 2000-08-04
;; PRIOR APPLICATION NUMBER: 60/223,378
;; PRIOR FILING DATE: 2000-08-07
;; NUMBER OF SEQ ID NOS: 9597
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 5724
;; LENGTH: 288
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (58)
;; OTHER INFORMATION: n=A,T,C or G
;; NAME/KEY: unsure
;; LOCATION: (59)
;; OTHER INFORMATION: n=A,T,C or G
US-09-796-692-5724

Query Match      45.3%; Score 272.4; DB 9; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 252 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGTTCAGTTCATGAGAAATTT- 310
DB 1 CCATCTTTATAGAAATATACAGTTTATCAAGATGAGTTCAGTTCATGAGAAATTTNC 60

QY 311 AGAAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTTCCATCAGTGGATTATAC 370
DB 61 AGAAGAAAGTTTCATCAGCTTGTATGACCGTGTGAGTTTCCATCAGTGGATTATAC 120

QY 371 TTGACCGGAATGTGTATTCAGGCTGTAAATGATGACAGAGATGCTGCACCAATC 430
DB 121 TTGACCGGAATGTGTATTCAGGCTGTAAATGATGACAGAGATGCTGCACCAATC 180

QY 431 ATTACGGGCACTCTCAGTGCAGATGACAGAGAGGTTTAAATGATGCTTGTATCAATTT 490
DB 181 ATTACGGGCACTCTCAGTGCAGATGACAGAGAGGTTTAAATGATGCTTGTATCAATTT 240

QY 491 TCAGATTGTGAATTTTGGCTGCTGTATATCTTTTGGAAATTTT 538
DB 241 TCAGATTGTGAATTTTGGCTGCTGTATATCTTTTGGAAATTTT 288

RESULT 24
US-10-040-862-5724
;; Sequence 5724, Application US/10040862
;; Publication No. US20030078396A1
;; GENERAL INFORMATION:
;; APPLICANT: Gaiger, Alexander
;; APPLICANT: Algate, Paul A.
;; APPLICANT: Mannion, Jane
;; APPLICANT: Retter, Marc
;; APPLICANT: Corixa Corporation
;; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
;; TITLE OF INVENTION: Hematological Malignancies
;; FILE REFERENCE: 014058-01352005
;; CURRENT APPLICATION NUMBER: US/10/040,862
;; PRIOR FILING DATE: 2001-11-06
;; PRIOR APPLICATION NUMBER: US 60/186,126
;; PRIOR FILING DATE: 2000-03-01
;; PRIOR APPLICATION NUMBER: US 60/190,479
```

```
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 60/223,416
/ PRIOR FILING DATE: 2000-08-04
/ PRIOR APPLICATION NUMBER: US 60/223,378
/ PRIOR FILING DATE: 2000-08-07
/ PRIOR APPLICATION NUMBER: US 09/796,692
/ PRIOR FILING DATE: 2001-03-01
/ NUMBER OF SEQ ID NOS: 10467
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 5724
/ LENGTH: 288
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (58)
/ OTHER INFORMATION: n=A,T,C or G
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (59)
/ OTHER INFORMATION: n=A,T,C or G
/ US-10-040-862-5724
```

```
Query Match      45.3%; Score 272.4; DB 14; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 252 CCATTCTTTATAGAAATTAATCAAGTTTAATCAAGATGAGTATGATGAGAAATTT- 310
Db 1 CCATTCTTTATAGAAATTAATCAAGTTTAATCAAGATGAGTATGATGAGAAATNNC 60

QY 311 AAGAGAAAGTTTCATAGCTTGTCTATGACCGTGCTAGTTTCCATGAGTGAATATACC 370
Db 61 AAGAGAAAGTTTCATAGCTTGTCTATGACCGTGCTAGTTTCCATGAGTGAATATACC 120

QY 371 TTGACCGGAATGTGTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATC 430
Db 121 TTGACCGGAATGTGTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATC 180

QY 431 ATTGACCGGCACTCAGTCCAGATGACATGACGAGGTTAATATGCTTTGATCATTTT 490
Db 181 ATTGACCGGCACTCAGTCCAGATGACATGACGAGGTTAATATGCTTTGATCATTTT 240

QY 491 TCAGATTGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTT 538
Db 241 TCAGATTGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTT 288
```

```
RESULT 25
US-10-057-475B-5724
/ Sequence 5724, Application US/10057475B
/ Publication No. US20040002068A1
/ GENERAL INFORMATION:
/ APPLICANT: Galger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mammion, Jane
/ APPLICANT: Clapper, Jonathan David
```

```
/ APPLICANT: Wang, Aljun
/ APPLICANT: Ordenez, Nadia
/ APPLICANT: Carrez, Lauren
/ APPLICANT: McNeill, Patricia Dianne
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ TITLE OF INVENTION: Hematological Malignancies
/ FILE REFERENCE: 014058-014402US
/ CURRENT APPLICATION NUMBER: US/10/057,475B
/ CURRENT FILING DATE: 2002-01-22
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: US 09/796,692
/ PRIOR FILING DATE: 2001-03-01
/ NUMBER OF SEQ ID NOS: 10979
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO: 5724
/ LENGTH: 288
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1)-(288)
/ OTHER INFORMATION: n = G, A, C or T
/ US-10-057-475B-5724
```

```
Query Match      45.3%; Score 272.4; DB 17; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 252 CCATTCTTTATAGAAATTAATCAAGTTTAATCAAGATGAGTATGATGAGAAATTT- 310
Db 1 CCATTCTTTATAGAAATTAATCAAGTTTAATCAAGATGAGTATGATGAGAAATNNC 60

QY 311 AAGAGAAAGTTTCATAGCTTGTCTATGACCGTGCTAGTTTCCATGAGTGAATATACC 370
Db 61 AAGAGAAAGTTTCATAGCTTGTCTATGACCGTGCTAGTTTCCATGAGTGAATATACC 120

QY 371 TTGACCGGAATGTGTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATC 430
Db 121 TTGACCGGAATGTGTATCCAGGCTGTAAATGAATGACAGAGATGCTGCACCAATC 180

QY 431 ATTGACCGGCACTCAGTCCAGATGACATGACGAGGTTAATATGCTTTGATCATTTT 490
Db 181 ATTGACCGGCACTCAGTCCAGATGACATGACGAGGTTAATATGCTTTGATCATTTT 240

QY 491 TCAGATTGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTT 538
Db 241 TCAGATTGTGAATTTTGGCTGCTGTATTAATCCTTTTGGGAATTTT 288
```

```
RESULT 26
US-10-154-884B-5724
/ Sequence 5724, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:
```



```
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc W.
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013521US
CURRENT APPLICATION NUMBER: US/10/154,884B
CURRENT FILING DATE: 2002-05-23
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 11290
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5724
LENGTH: 288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(288)
OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-5724

Query Match      45.3%; Score 272.4; DB 17; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 252 CCATCTTTATAGGAATATCAGTTTAATCAAGTAGAGTGAATTGATGAGAAATTT- 310
DB 1 CCATCTTTATAGGAATATCAGTTTAATCAAGTAGAGTGAATTGATGAGAAATNNC 60
QY 311 AAGAAGAAAGTTATCAGTTGCTATGACCGTGTCAAGTTTCATCAGTGAATTATAC 370
DB 61 AAGAAGAAAGTTATCAGTTGCTATGACCGTGTCAAGTTTCATCAGTGAATTATAC 120
QY 371 TTTCACCGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCCAATC 430
DB 121 TTTCACCGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCCAATC 180
QY 431 ATTACGCGCACTCACTGCGCAAGTCAATGACGGGTTAAATGCTTTGATCAATTT 490
DB 181 ATTACGCGCACTCACTGCGCAAGTCAATGACGGGTTAAATGCTTTGATCAATTT 240
QY 491 TCAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTTT 538
DB 241 TCAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTTT 288
```

```
RESULT 27
US-10-764-324-5724
; Sequence 5724, Application US/10764324
; Publication No. US20040175739A1
; GENERAL INFORMATION:
```

```
APPLICANT: Gaiger, Alexander
APPLICANT: Algate, Paul A.
APPLICANT: Mannion, Jane
APPLICANT: Retter, Marc
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
FILE REFERENCE: 014058-013520US
CURRENT APPLICATION NUMBER: US/10/764,324
CURRENT FILING DATE: 2004-01-23
PRIOR APPLICATION NUMBER: US/10/040,862
PRIOR FILING DATE: 2001-11-06
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10467
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 5724
LENGTH: 288
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: unsure
LOCATION: (58)
OTHER INFORMATION: n=A,T,C or G
NAME/KEY: unsure
LOCATION: (59)
OTHER INFORMATION: n=A,T,C or G
US-10-764-324-5724
```

```
Query Match      45.3%; Score 272.4; DB 19; Length 288;
Best Local Similarity 98.6%; Pred. No. 5.4e-71;
Matches 284; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 252 CCATCTTTATAGGAATATCAGTTTAATCAAGTAGAGTGAATTGATGAGAAATTT- 310
DB 1 CCATCTTTATAGGAATATCAGTTTAATCAAGTAGAGTGAATTGATGAGAAATNNC 60
QY 311 AAGAAGAAAGTTATCAGTTGCTATGACCGTGTCAAGTTTCATCAGTGAATTATAC 370
DB 61 AAGAAGAAAGTTATCAGTTGCTATGACCGTGTCAAGTTTCATCAGTGAATTATAC 120
QY 371 TTTCACCGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCCAATC 430
DB 121 TTTCACCGGAATGTGTTATCCAGGCTGTTAAATGAATGACAGAGATGCTGCCAATC 180
QY 431 ATTACGCGCACTCACTGCGCAAGTCAATGACGGGTTAAATGCTTTGATCAATTT 490
DB 181 ATTACGCGCACTCACTGCGCAAGTCAATGACGGGTTAAATGCTTTGATCAATTT 240
QY 491 TCAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTTT 538
DB 241 TCAGATTGGAATTTTGGCTGCTGTATATATCCTTTGGGAATTTT 288
```

```
RESULT 28
```

```

US-10-029-386-25411
; Sequence 25411, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25411
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC012678.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.51
; OTHER INFORMATION: SWISSPROT HIT: Q61768, EVALUE 1.10e-01
; OTHER INFORMATION: NT HIT: AF12095.1, EVALUE 1.30e-01
; OTHER INFORMATION: EST_HUMAN HIT: BF744133.1, EVALUE 2.00e-59
US-10-029-386-25411

Query Match      39.5%; Score 237.4; DB 16; Length 527;
Best Local Similarity 65.7%; Pred. No. 2.5e-60;
Matches 346; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 75 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 134
DB 1 AGCCCAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 60
QY 135 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 194
DB 61 ATGACACAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 120
QY 195 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 254
DB 121 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 180
QY 255 TTCTTATAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 314
DB 181 TCTCTACAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 240
QY 315 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 374
DB 241 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 300
QY 375 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 434
DB 301 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 360
QY 435 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 494
DB 361 AGGCAAAAGAGAGATCTTGAGTAATGGTTCCTCAATCCATGCGACCACTTAATAG 420
QY 495 ATTGTAATTTTGGTTCCTTGATTAATCTTTTGGGAAATTTTAAACCCCACTTAATAG 554
DB 421 ATTGTAATTTTGGTTCCTTGATTAATCTTTTGGGAAATTTTAAACCCCACTTAATAG 480
QY 555 AACTATGTATGTATCAACAAATGTGATGAAGAAATATATGA 601
DB 481 GATTTGTGAAGAAATCAATTAATGTTCTATGATGAAGAAATGCTTTTAA 527

RESULT 29
US-10-094-466-41
; Sequence 41, Application US/10094466
; Publication No. US20030203363A1
; GENERAL INFORMATION:
; APPLICANT: Spylek et al.

```

```

; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
; TITLE OF INVENTION: AND METHODS OF USING
; TITLE OF INVENTION: THE SAME
; FILE REFERENCE: 21402-290D
; CURRENT APPLICATION NUMBER: US/10/094,466
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/288,148
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/338,375
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 60/275,579
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/335,302
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/275,601
; PRIOR FILING DATE: 2001-03-13
; PRIOR APPLICATION NUMBER: 60/276,000
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/277,338
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patin 2.1
; SEQ ID NO 41
; LENGTH: 645
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (26) .. (584)
US-10-094-466-41

Query Match      34.6%; Score 207.8; DB 17; Length 645;
Best Local Similarity 59.9%; Pred. No. 2.2e-51;
Matches 347; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 23 CCGACTCTCCGATGGCCACAGATGCTTAAATCCAAAACTGGCGTTCAGCACAA 82
DB 8 CTGTGTCCCGCCAGCCATGACACCTTCAGACCAAGAGCTGCTGAGGCGGAG 67
QY 83 AAGAAGATCTTGGTAAATGGTTCCTCAATCCATGCGACCACTTAATAGAGACAA 142
DB 68 AAGAAGATCTTGGTAAATGGTTCCTCAATCCATGCGACCACTTAATAGAGACAA 127
QY 143 AGTAGAGAGTCTGATGAGTCTTACAGAGTGAACAGGAGTAAACCCAAAGAAAG 202
DB 128 AGTAGAGAGTCTGATGAGTCTTACAGAGTGAACAGGAGTAAACCCAAAGAAAG 187
QY 203 GAGGCGAGAGATCATCAAGAACCTCATCAGACAGTCAAGCTTGAAGCTTTTAT 262
DB 188 GAGGCGAGAGATCATCAAGAACCTCATCAGACAGTCAAGCTTGAAGCTTTTAT 247
QY 263 AGGAATTAATCAATTAATCAAGATGAGTGAAGAAATTTAAGAAAGAAATTT 322
DB 248 CGTAGGAGACCAAGCTGGGCGGTGAGAGGCTGGGCTGGGCTGGGCTGGGCTGG 307
QY 323 CATCAGCTTGTATGACGCTGCTGCTTCAATCAGTGTGATTAATCTTTGACCGGAT 382
DB 308 CCGTGTGGGCTGATGACGCTGCTGCTTCAATCAGTGTGATTAATCTTTGACCGGAT 367
QY 383 GTTTATCAGGCTGTTTAAATGAATGACAGAGATGCTGACCAATCATTTGAGCGCCAC 442
DB 368 GTTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 427
QY 443 CTCACCTGCAAGTCAATGACGAGGTTTAAATGCTTTGATCAATTTTTCAGATTGTGA 502
DB 428 CTGACGCGCAATGCTCAAGCGGCTGATCAACAGGCTGTTGGGCTGCTGCTGCTGCTG 487

```

QY 503 TTTTGGCTGCTGTATATCTTTGGGAATTTTAAACCCACTTACAAAAACTATGT 562
Db 488 TTCTTGCTGCTGCTGTACGGCCCGCCGAGCCCTACCGCTCCCACTGCGAGGATCTGC 547
QY 563 GATGATATCAAAATGTTGATGAGAGACATATGA 601
Db 548 GAGGCGCTGGCGGATGCTGAGACGAGGCGACCTCTGA 586

RESULT 30

US-10-416-314-70
; Sequence 70, Application US/10416314
; Publication No. US20040082308A1
; GENERAL INFORMATION:

APPLICANT: YUE, Henry
APPLICANT: GANDHI, Ameena R.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: SHARNAKAR, Anita
APPLICANT: CHAWLA, Narinder K.
APPLICANT: SANJANWALA, Madhusudan M.
APPLICANT: THORNTON, Michael B.
APPLICANT: ELLIOTT, Vicki S.
APPLICANT: LU, Yan
APPLICANT: GIETZEN, Kimberly J.
APPLICANT: BURGARD, Neil
APPLICANT: DING, Li
APPLICANT: HAPALIA, April J.A.
APPLICANT: TANG, Y. Tom
APPLICANT: BANDMAN, Olga
APPLICANT: WARREN, Bridget A.
APPLICANT: HONCHELL, Cynthia D.
APPLICANT: LU, Dying Aina M.
APPLICANT: THANGAVELU, Kavitha
APPLICANT: LEE, Sally
APPLICANT: XU, Yuming
APPLICANT: YANG, Yuming
APPLICANT: LAL, Preeti G.
APPLICANT: TRAN, Bao
APPLICANT: ISON, Craig H.
APPLICANT: DUGGAN, Brendan M.
APPLICANT: KAREHT, Stephanie K.
TITLE OF INVENTION: SECRETED PROTEINS
FILE REFERENCE: PI-0287 USN
CURRENT APPLICATION NUMBER: US/10/416,314
CURRENT FILING DATE: 2003-05-08
PRIOR APPLICATION NUMBER: US 60/247,505
PRIOR FILING DATE: 2000-11-08
PRIOR APPLICATION NUMBER: US 60/249,642
PRIOR FILING DATE: 2000-11-09
PRIOR APPLICATION NUMBER: US 60/249,824
PRIOR FILING DATE: 2000-11-16
PRIOR APPLICATION NUMBER: US 60/252,824
PRIOR FILING DATE: 2000-11-21
PRIOR APPLICATION NUMBER: US 60/254,305
PRIOR FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: US 60/256,448
PRIOR FILING DATE: 2000-12-18
NUMBER OF SEQ ID NOS: 130
SOFTWARE: PERL Program
SEQ ID NO 70
LENGTH: 2108
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incycle ID No: 2013095CBI
US-10-416-314-70

Query Match 34.6%; Score 207.8; DB 18; Length 2108;
Best Local Similarity 59.9%; Pred. No. 4,4e-51;
Matches 347; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

QY 23 CCGACTCTCCGATGCGCACAGATGTCCTTAATTCAAAACTGGCCGTTGACGGACAA 82
Db 412 CTGTGTCCCGCCGAGGCCATGACACCTTCAGACCAAGAGCTGGCTGACAGGCGCAG 471
QY 83 AAGAAGATCTTGGTAAATGATGTCGAATTCATGCGCCACCTTAATAGACGACAA 142
Db 472 AAGAAGCTCTGTAGTAAGATGCGCTCAAGCGAGCGATGCGCTGCTGAGATGACCC 531
QY 143 AGTAGTAGGTCGTGATGAGCTCTACAGATGACCGAGGAGATACCCAAAAAAGAG 202
Db 532 AGCAGTAGAGTGTGTGATGAGCTGTACCGGCCACCAAGAGATTCACGCCACCGCAG 591
QY 203 GAGGCAGAGAAAGATCATCAAGAACTTCATCAAGACGTCAAGCTGGCCATTTTAT 262
Db 592 GAGGCCAGAAAGATGCTCAAGAACTGTCTCAAGGTGGCCCTGAAGCTGGAGACTGTGCTG 651
QY 263 AGGAATATCATGTTTATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGTT 322
Db 652 CGTGGGACCAAGCTGGGCGGTGAGAGCTGGCGCTGCTGCGGCTTCGCCACCGGCG 711
QY 323 CATCAGCTTGTGATGACCGGTGATGATTCATCAGGTGATTAATCTTTGACCGGAT 382
Db 712 CGCTGCTGTGCTGATGACCGGCTGACGCTTCCACCAAGTGAATTCACCTTCGACCGGCG 771
QY 383 GTGTTATCCAGGCTGTAAATGAAATGACAGAGATGCTGACCAATTCATTCAGCGGAC 442
Db 772 GTGCTGCGCGCGGCTGTGCGAGTGCGCGGACCTGCTGACCAAGCGCGTGGCTCCCGAC 831
QY 443 CTCACGTGCCAATGATGATGAGCGGTTAATATGCTTTGATATTTTCAATGTTGAA 502
Db 832 CTGACCGCAAGTCCACCGCGCCATCAACCAAGTGTGCGCCACCTGACCGGCTGCGAC 891
QY 503 TTTTGGCTGCTGTATATCTTTGGGAATTTTAAACCCACTTACAAAAACTATGT 562
Db 892 TTCTTGCTGCTGCTGTACGGCCCGCCGAGCCCTACCGCTCCACCTGCGAGATCTGC 951
QY 563 GATGATATCAAAATGTTGATGAGAGAAACATATGA 601
Db 952 GAGGCGCTGGCGGATGCTGAGCGAGGCGAGCTCTGA 990

RESULT 31

US-09-816-828-5
; Sequence 5, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Ma, Yundong
APPLICANT: Wang, Zhiwei
APPLICANT: Zhao, Qing A.
APPLICANT: Zhang, Jie
APPLICANT: Wang, Jian-Rui
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20020150898A1 Nucleic Acids and
FILE REFERENCE: Polypeptides
CURRENT APPLICATION NUMBER: US/09/816,828
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 09/552,929
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 5
LENGTH: 1188
TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (174)..(731)
US-09-816-828-5

Query Match 34.2%; Score 205.6; DB 9; Length 1188;
Best Local Similarity 60.0%; Pred. No. 1.4e-50;
Matches 343; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

```
QY 30 CTCGATGCGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTACAGGCACAAAGAGA 89
DB 163 CTGCTAAGCGCCATGACCTTACAGCACAAGAGCTGGCTTGACGCGCAGAAAGAAC 222
QY 90 TCTTGGGTAATATGCTGCTCAATTCATGGCCACCTTAATAGACACAAAGATGTC 149
DB 223 TCTGATGATGAGTGGGCTCAAGGAGTGGCTGCTGCTGATGACACAGAGTGT 282
QY 150 AGGTGCTGATGAGCTCTACAGAGTACAGGAGTACACCAAAACAAGAGAGGACAG 209
DB 283 AGGTGCTGATGAGCTGTCACGCGCACAGGAGTTCACGCGCAGCGCAGAGAGGCC 342
QY 210 AGAAGATCATCAAGAACTTCATCAAGACATCAAGCTGGCCATTTTATAGAAATA 269
DB 343 AGAAGATGCTCAAGAACTTCATCAAGATGAGCTGGAAGCTGCTGCTGCGGG 402
QY 270 ATGATGTTAATCAAGATGAGCTGATGATGAGAAATTAAGAAAGTTATCATCAGC 329
DB 403 ACCAGCTGGGCGGTGAGAGCTGGCTGCTGCGGCTTCGCCACCGGGCGGCTGCC 462
QY 330 TTGCTATGACCGTGTGCTGCTTTCATCAGGTGATTAATATTCAGCCGATGTGTAT 389
DB 463 TGGCCATGAGCGGCGTCAAGCTTCAACAGGTGATCTTACCTTCGACCGGCGGCTGCT 522
QY 390 CCAAGCTGTTAATATGATGACAGAGTGTGACCAATCATTCAGCGCCACTCTACG 449
DB 523 CGGTGGGCTGCTGAGTCCGCGGACCTGCTGACCAAGGCGGTGCCACTCTACG 582
QY 450 CCAAGTCATGAGCGGCTTAATATGCTGATGATTTTTCAGATTTGATTTTGG 509
DB 583 CCAAGTCACCGGCGGCTCAACAGAGTGTGCGGCACTGAGCGACTGCTCTGG 642
QY 510 CTGCTTGTATATCTTTTGGGAATTTTAAACCCCACTTACAAATATATGATGATGA 569
DB 643 CTGCGCTTACAGGCGCGCGGAGCCCTACCGCTCCACCTGCGCAGAGATGTGAGAGGCC 702
QY 570 TCAACAAAATGTTGATGAGAGACATATGA 601
DB 703 TGGGCGGATGCTGAGAGGAGGCGCTCTGA 734
```

RESULT 32
US-10-094-466-43
Sequence 43, Application US/10094466
Publication No. US2003020363A1
GENERAL INFORMATION:
APPLICANT: Splytek et al
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM
TITLE OF INVENTION: AND METHODS OF USING
FILE REFERENCE: 21402-290D
CURRENT APPLICATION NUMBER: US/10/094,466
PRIOR FILING DATE: 2002-03-07
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/288,148
PRIOR FILING DATE: 2001-05-02
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/338,375
PRIOR FILING DATE: 2001-12-04

PRIOR APPLICATION NUMBER: 60/275,579
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/335,302
PRIOR FILING DATE: 2001-10-31
PRIOR APPLICATION NUMBER: 60/275,601
PRIOR FILING DATE: 2001-03-13
PRIOR APPLICATION NUMBER: 60/276,000
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/277,338
PRIOR FILING DATE: 2001-03-20
Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 114
SOFTWARE: Patin 2.1
SEQ ID NO 43
LENGTH: 619
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(566)
US-10-094-466-43

Query Match 33.8%; Score 203.4; DB 17; Length 619;
Best Local Similarity 60.0%; Pred. No. 4.4e-50;
Matches 339; Conservative 0; Mismatches 226; Indels 0; Gaps 0;

```
QY 37 GGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTACAGGCACAAAGAAATCTGGG 96
DB 4 GGCCATGACACCTTCAGACCAAGAGCTGCTGCTGACGGGCGCAAGAAAGCTCTTG 63
QY 97 TAAATGCTGTCACCAATCCATGCGCACCTTAATAGACGACACAGATGATGAGTGTCT 156
DB 64 TAAATGCTGTCACCAAGGAGCTGCTGCTGCTGATGATGACACAGAGTGAAGTGTCT 123
QY 157 GGATGACTCTACAGATGACCAAGGAGTACACCCAAACAAAGAGAGCGACAGAAAT 216
DB 124 GGATGAGCTGATCCGCGCACCAAGGAGTTCACGGGCGAGCCCAAGAGAGGCGCAGAAAT 183
QY 217 CATCAAGAACTTCATCAAGACAGTCAAGCTGCGCATCTTTATAGGAATATCATGTT 276
DB 184 GCTCAAGAACTGCTCAAGAGTGGCCCTGAGAGTGGAGCTGCTGCTGCGGAGCCAGCT 243
QY 277 TAATCAAGATGAGCTAGCATTTGATGAGAGAAATTTAAGAAAGTTTCACTGCTAT 336
DB 244 GGGCGGTAGAGAGCTGGGCTGCTGCTGCGGCGCACCGGGGCGCTGCTGCGCAT 303
QY 337 GACCGTGTGCTTTCATCAGGTGATTAATCTTTTGAACCGGAAATGTGTTATCAAGCT 396
DB 304 GACGCGCTGACCTTCCACCAAGTGAATTCATCTTTCGACCGGCGGCTGCGCGCG 363
QY 397 GTTAAATGAGACAGAGATGCTGACCAATATTCAGGCGGCACTGAGCCAGATC 456
DB 364 GCTGCTGAGTCCCGGAGCTGCTGACAGGCGGCTGCTGCCACCTGAGCGCAAGTC 423
QY 457 ACATGAGCGGTTAATATGCTTTGATGATTTTCAATTTTCAATTTTGGTGCCTT 516
DB 424 CCAAGCGGCTCAACACAGTGTGGGCACTTACCGACCTGCACTTCTGCTGCGCT 483
QY 517 GTATATCTTTTGGGAATTTTAAACCCCACTTACAAAAACTATGATGATGATCAACA 576
DB 484 CTACGCGCCCGCGAGGCTTACCGCTCCACTGCGCAGAGATCTGAGAGGCGCTGGCGG 543
QY 577 AATGTTGATGAGAGACATATGA 601
DB 544 GATGCTGACGAGGCGAGCTCTGA 568
```

RESULT 33
US-10-764-420-2116
Sequence 2116, Application US/10764420
Publication No. US20050084872A1
GENERAL INFORMATION:
APPLICANT: Lum, Pek Yee

```
; APPLICANT: Tan, Yejun
; APPLICANT: Dai, Hongyue
; TITLE OF INVENTION: Methods For Determining Whether An Agent
; TITLE OF INVENTION: Possesses A Defined Biological Activity
; FILE REFERENCE: ROSA122057
; CURRENT APPLICATION NUMBER: US/10/764.420
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: US 60/442,797
; PRIOR FILING DATE: 2003-01-24
; PRIOR APPLICATION NUMBER: US 60/474,413
; PRIOR FILING DATE: 2003-05-30
; NUMBER OF SEQ ID NOS: 3683
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2116
; LENGTH: 1389
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 80
; OTHER INFORMATION: n = A,T,C or G
US-10-764-420-2116
```

```
Query Match          33.2%; Score 199.8; DB 21; Length 1389;
Best Local Similarity 59.7%; Pred. No. 8.6e-49;
Matches 336; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

QY 39 CCACAGATGCTTTTAATTCACAAAACCTGGCCGTTACAGGACAAAGAGATCTTGAGTA 98
DB 82 CCATGGACACCTTCAGACGAGAGAGCTGGCCCTGACAGGCCCAAGAAAGTCTTGACA 141
QY 99 AATGCTGTCCAAATTCATGCCCAACCTTAATAGACACAAAGATGAGGTGCTGG 158
DB 142 AGATGGCTTCCAGAGCATGCTGCTGTGTGTGCAATATCCAGACAGAGAGTCTTGG 201
QY 159 ATGAGCTCTACAGAGTGACACGAGGATACACCCAAAACAGAGAGAGAGAGATCA 218
DB 202 ATGATAGTACAGAGGACCAAGAGAGTTCACGCGACGCGAGAGAGAGAGAGAGTAG 261
QY 219 TCAAGAACCTATCAGACAGATCATCAAGCTGGCCATCTTTATAGGAATATCAGTTTA 278
DB 262 TGAAGAACCTGTGAAGAGTGGCTGTGAAGCTGGCTGTGTGTGTGTGTGTGTGTGTGT 321
QY 279 ATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGAGTTCATCAGCTTGCTATGA 338
DB 322 ACAGCATGAGCTGGCCAGCTGACAGCGGTTCCGGGGCCCGGTCCGCAACCTGGCCATGA 381
QY 339 CCGGTCAGTTCATCAGAGTGAATTAATCTTTGACCGGAATGTGTTATCCAGGCTGT 398
DB 382 CAGCCCTCAGCTTCACAGAGTGAATCTTCACTTTGACCGCGCTGTGCTGGCCACTGGGC 441
QY 399 TAAATGATCAGAGAGATCTGACCAAAATCATTTAGCGCCACTCAGTCCCAAGTCAAC 458
DB 442 TGTGAGAGTCAAGAGAGCTGTGTGACCAAGCTATGTGCCCGGCACTTACAGTCC 501
QY 459 ATGAGCGGATTAATATGCTTTGATCATTTTTCAGATGTGAATTTTGGCTGCTGT 518
DB 502 ACCGCGCATCAATCAATCTTCACTTGTGCAATGTGATGATCTTCTGGCGCGCTGT 561
QY 519 ATATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATGATCAACAAA 578
DB 562 ACAGCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 621
QY 579 TGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
DB 622 TGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 644
```

```
RESULT 34
US-09-822-849A-81
; Sequence 81, Application US/09822849A
; Patent No. US20020045170A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fehel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulubeta, Kamalakara
; APPLICANT: Graham, James R.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6403
; CURRENT APPLICATION NUMBER: US/09/822.849A
; PRIOR APPLICATION NUMBER: 60/195,582
; PRIOR FILING DATE: 2001-09-04
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 598
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 81
; LENGTH: 1113
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-822-849A-81
```

```
Query Match          30.3%; Score 182; DB 9; Length 1113;
Best Local Similarity 58.7%; Pred. No. 1.7e-43;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGATGGCCACAGATGCTTTAATTCACAAAACCTGGCCGTTACAGGACAAAGAGAGA 89
DB 52 CTCGAGAGACCATGAGAGTCTTTCATGCTCAAAAGAGCTGGACATCGACAGAGAGAGAG 111
QY 90 TCTTGGTAAATGTGTCTCAAAATTCATGCGCCACACTTAATAGACAGACAGATGTG 149
DB 112 TACTGATGAATAGAGGCGGCTGCTGTGTGCTATCTTCAATAGATGAGACAAAGCAGTG 171
QY 150 AGGTGCTGATGAGCTCTACAGAGTACCGAGAGATACCCAAAACAGAGAGAGAGAG 209
DB 172 AGGTGCTGATGAGCTCTACAGAGTACCGAGAGATACCGAGAGAGAGAGAGAGAGAG 231
QY 210 AGAAGATCATCAAGACCTCATCAAGACAGTGCATCAAGCTGGCCATCTTTATAGAGATA 269
DB 232 AGGCGCTGATCAAGAGACTGATCAAGAGTGCATCAAGAGTGTGTGTGTGTGTGTGTGT 291
QY 270 ATCAGTTAATCAAGATGAGCTAGCTTGAATGAGAGAAATTTAAGAAAGAGTTCATCAGC 329
DB 292 GCTCTTTGGCCCAAGTGAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 351
QY 330 TTGCTATGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 389
DB 352 GTGCCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 411
QY 390 CCGAGCTGTTAATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 449
DB 412 CTGGCTGCTGACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 471
QY 450 CCAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 509
DB 472 CCAAGTCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 531
QY 510 CTGCTCTGTATATCTCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATG 569
DB 532 CGGCGCTCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 585
QY 570 TCAACAAAATGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
DB 586 TCAAGAGAGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 617
```

```
RESULT 35
US-10-363-374-15
; Sequence 15, Application US/10363374
; Publication No. US20050048483A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Su, Eric
; APPLICANT: Mang, He
; TITLE OF INVENTION: NOVEL SECRETED PROTEINS AND THEIR USES
; FILE REFERENCE: X13020
; CURRENT APPLICATION NUMBER: US/10/363,374
; CURRENT FILING DATE: 2003-02-27
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 1165
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (100)..(651)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (100)..(180)
; US-10-363-374-15

```

```

Query Match      30.3%; Score 182; DB 21; Length 1165;
Best Local Similarity 58.7%; Pred. No. 1.7e-43;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

```

```

QY 30 CTCGATGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTCAAGCACAAGAAAGA 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 89 CTCGAGAGACCCATGAGAGTCTTCAAGCTCAAGAGCCTGGCACTGCAAGCAGAGAAAGC 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 90 TCTTGGGTAAATGATGTGTCAAATTCATGCGCACCATTTAATAGACAGACAAATAGTG 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 149 TACTGAGTAAGATGGCGGGTGGCTGTGTGGCTCATCTTTCATAGATGAGACAGCAGTG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 150 AGGTGCTGATGAGCTCTACAGAGTGAACGAGAGATACACCAACAAAGAGAGCAG 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 209 AGGTGCTAAGATGAGCTCTACCGGTGTCCAAAGAGTACAGCAGACCGGCCCGCCAGGCC 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 210 AGAAGATCATCAAGAACTCATCAAGACAGTCAAGCTGAGCCATTTTATAGAAATA 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 269 AGGCGCTGATCAAGAGCTGATCAAGAGTGCATCAAGGTGGCTGTGCTCACCGCAATG 328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 270 ATCAGTTTAATCAAGATGAGCTAGATGATGAGAAATTTAAGAAAGATTTCATCAGC 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 329 GCTCTTTGGCCCAAGTGAAGTGGCCCTGGCTACCGCTTTCGCCAAGAGCTGGCAGG 388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 330 TTGCTATGACCGTGTCAAGTTCCATCAGGTGATATACCTTTGACCGGAATGTGTAT 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 389 GTGCCATGACGGCATTAAGCTTTGTGAGGTAGACTTCACCTTCGAGGCTGTCTGTG 448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 390 CCAAGCTGTTAATGATGAGAGATGCTGCAACCAATCATTCAGCGCACCTCAGTG 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 449 CTGGCTGTGACCGAGTGGCGGGATGTGCTGCTAGAGTTGTGGAACACACCTCAGCG 508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 450 CCAAGTCAATGAGCGGGTTAATATATGTCTTTGATCATTTTTCAGATTGTGAATTTTGG 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 509 CCAAGTCAATGAGCGGCATCCGACAGTGTGATCATCTCTGACCCAGGTCTGTCTCA 568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 510 CTGGCTGTGATATCTTTTGGGAATTTTAAACCCCATTAACAAAGATATGTGATGTA 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 569 CGGCTCTATATGAGGCTCT-----GACTTCACTCAGACCTTGGCAAGATCTGTGAGCGAC 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 570 TCACCAAAATGTTGATGAGAGACATATGA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 623 TCAGGAAGCTGTAGACGAAGGGAAGCTCTGA 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 36

```

US-10-302-172-537
; Sequence 537, Application US/10302172
; Publication No. US20040053250A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Drmanac, Radoje T.

```

```

; TITLE OF INVENTION: No. US20040053250A1 Arginine-rich Protein-like Nucleic Acids an
; APPLICANT: Polypeptides
; FILE REFERENCE: 803_1CNCP
; CURRENT APPLICATION NUMBER: US/10/302,172
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/225,251
; PRIOR FILING DATE: 2002-08-20
; PRIOR APPLICATION NUMBER: PCT US02/05095
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 09/799,451
; PRIOR FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 950
; SOFTWARE: pc FL_genes Version 2.0
; SEQ ID NO 537
; LENGTH: 1175
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (113)..(664)
; US-10-302-172-537

```

```

Query Match      30.3%; Score 182; DB 18; Length 1175;
Best Local Similarity 58.7%; Pred. No. 1.7e-43;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

```

```

QY 30 CTCGATGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTCAAGCACAAGAAAGA 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 102 CTCGAGAGACCCATGAGAGTCTTCAAGCTCAAGAGCCTGGCACTGCAAGCAGAGAAAGC 161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 90 TCTTGGGTAAATGATGTGTCAAATTCATGCGCACCATTTAATAGACAGACAAATAGTG 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 162 TACTGAGTAAGATGGCGGGTGGCTGTGTGTCTCATCTTTCATAGATGAGACAGCAGTG 221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 150 AGGTGCTGATGAGCTCTACAGAGTGAACGAGAGATACACCAACAAAGAGAGCAG 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 222 AGGTGCTAAGATGAGCTCTACCGGTGTCCAAAGAGTACAGCAGACCGGCCCGCCAGGCC 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 210 AGAAGATCATCAAGAACTCATCAAGACAGTATCAAGCTGAGCCATTTTATAGAAATA 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 282 AGGCGGTGATCAAGAGCTGATCAAGAGTGCATCAAGGTGGCTGTGCTGACCGCAATG 341
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 270 ATCAGTTTAATCAAGATGAGCTAGATGATGAGAAATTTAAGAAAGATTTCATCAGC 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 342 GCTCTTTGGCCCAAGTGAAGTGGCCCTGGCTACCGCTTTCGCCAAGAGCTGGCAGG 401
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 330 TTGCTATGACCGTGTCAAGTTCCATCAGGTGATATACCTTTGACCGGAATGTGTAT 389
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 402 GTGCCATGACGGCATTAAGCTTTGTGAGGTAGACTTCACCTTCAGAGGCTGTCTGTG 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 390 CCAAGCTGTTAATGATGAGAGATGCTGCAACCAATCATTCAGCGCCACCTCAGTG 449
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 462 CTGGCTGTGACCGAGTGGCGGGATGTGCTGTGAGTGTGGAACCAACCTCAGCGC 521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 450 CCAAGTCAATGAGCGGGTTAATATATGTCTTTGATCATTTTTCAGATTGTGAATTTTGG 509
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 522 CCAAGTCAATGAGCGGCATCCGACAGTGTGATCATCTCTGACCCAGGTCTGTCTCA 581
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 510 CTGGCTGTGATATCTTTTGGGAATTTTAAACCCCATTAACAAAGATATGTGATGTA 569
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 582 CGGCTCTATATGAGGCTCT-----GACTTCACTCAGACCTTGGCAAGATCTGTGAGCGAC 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 570 TCACCAAAATGTTGATGAGAGACATATGA 601
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 636 TCAGGAAGCTGTAGACGAAGGGAAGCTCTGA 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 37

```

US-10-959-539-59
; Sequence 59, Application US/10959539
; Publication No. US20050048623A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.

```

APPLICANT: HILLMAN, Jennifer L.
APPLICANT: TAL, Preeti
APPLICANT: TANG, Y. Tom
APPLICANT: YOE, Henry
APPLICANT: AU-YOUNG, Janice
APPLICANT: BANDMAN, Olga
APPLICANT: AZIMZAI, Yalda
APPLICANT: YANG, Junming
APPLICANT: LU, Dying Aina M.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: PATTERSON, Chandra
APPLICANT: SHAH, Puri
TITLE OF INVENTION: CELL CYCLE AND PROLIFERATION PROTEINS
FILE REFERENCE: PR-0722 PCT
CURRENT APPLICATION NUMBER: US/10/959,539
PRIOR FILING DATE: 2004-10-06
PRIOR APPLICATION NUMBER: US/10/031,915
PRIOR FILING DATE: 2002-01-18
PRIOR APPLICATION NUMBER: 60/145,075; 60/153,129; 60/164,647
PRIOR FILING DATE: 1999-07-21; 1999-09-08; 1999-11-10
NUMBER OF SEQ ID NOS: 108
SOFTWARE: PERL Program
SEQ ID NO 59
LENGTH: 1268
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 1558289CB1
US-10-959-539-59

Query Match 30.3%; Score 182; DB 21; Length 1268;
Best Local Similarity 58.7%; Pred. No. 1.8e-43;
Matches 336; Conservative 0; Mismatches 230; Indels 6; Gaps 1;

QY 30 CTCGAGGCGACAGATGCTTTAATTCAGAAACCTGGCCCTTCAGGCGACAAAGAG 89
DB 92 CTCAGAGACCCAGAGAGTCTTCAAGCTCAAGAGCCCTGGACCTGACAGAGAAAG 151
QY 90 TCTTGGGTAAGATGCTGTCCTCAATCCATGCGCCACCTTAATAGAGAGCAAGTAGT 149
DB 152 TACTGAGTAAGATGGGGGCGCTCTGTGCTCATCTCTTCATAGATGAGACAGAGT 211
QY 150 AGGTGCTGATGAGCTCTACAGAGTACAGAGGAGTACACCCAAACAGAGAGAGG 209
DB 212 AGGTGCTGATGAGCTCTACAGAGTACAGAGGAGTACAGAGGAGGAGGAGG 271
QY 210 AAGAGTATCAAGAACTTCATCAAGAGCTATCAAGCTGGCATTCTTTATAGAAAT 269
DB 272 AGGCGGTATCAAGAGCTGATCAAGGCTATCAAGGCTGTGTGCTGACCCGCAATG 331
QY 270 ATCAGTTATCAAGATGAGCTGATGATGAGAAATTTAAGAAAGTTATCATCAG 329
DB 332 GCTCCTTTGGCCCATGAGCTGCTGCTGCTACCCGCTTCCCGAAGCTCGGCGAG 391
QY 330 TTGCTATGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 389
DB 392 GTGCGCATGACGAGCTTACCTTTGGAGGTAGACTTCACTTCAGAGGCTGCTGCTG 451
QY 390 CCAAGGTATTAATGATGAGAGATGCTGACCAAAATCATTCAGCCGCTACTACTG 449
DB 452 CTGGCGCTGCTGACCGAGTGGCGGAGTGTCTGTAGAGTGGAGAACACACTCAGC 511
QY 450 CCAAGTACATGACCGGCTTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 509
DB 512 CCAAGTACATGACCGGCTTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 571
QY 510 CTGCTGTTATATCTTTTGGGAATTTAAACCCACTTACAAAACTATGTGATGTA 569
DB 572 CGGCGCTCTATGGGCT-----GACTTCACTCAGACCTTGGCAAGATCTGTGAGGAC 625
QY 570 TCAAGAAATGTTGATGAAGAGACATATGA 601

DB 626 TCAGGAAGCTGCTAGAGAGGAGGAGAGCTCTGA 657

RESULT 38
US-10-029-386-9211
Sequence 9211, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: AEOMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
PRIOR FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 9211
LENGTH: 500
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC035144.2
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
OTHER INFORMATION: NT HIT: g114725421, EVALUATE 0.00e+00
OTHER INFORMATION: EST_HUMAN HIT: B6178783.1, EVALUATE 0.00e+00
US-10-029-386-9211

Query Match 28.0%; Score 168.4; DB 16; Length 500;
Best Local Similarity 99.4%; Pred. No. 1.2e-39;
Matches 169; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 432 TTCAGGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 491
DB 1 TTCAGGCGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 492 CAGATTGTGAATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 551
DB 61 CAGATTGTGAATTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 552 AAAAATATGATGATGATCAAAATGTTGATGATGATGATGATGATGATGATGATG 601
DB 121 AAAAATATGATGATGATCAAAATGTTGATGATGATGATGATGATGATGATGATG 170

RESULT 39
US-09-918-995-35521

Sequence 35521, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: HySeq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 35521
LENGTH: 422
TYPE: DNA
ORGANISM: Homo sapiens
US-09-918-995-35521

Query Match 21.4%; Score 128.4; DB 10; Length 422;
Best Local Similarity 60.7%; Pred. No. 1.1e-27;

RESULT 41
US-10-191-803-330/c
; Sequence 330, Application US/10191803

RESULT 42
US-09-796-692-4917/C
; Sequence 4917, Application US/09796692
; Publication No. US20020198362A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; TITLE OR INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, D
; TITLE OF INVENTION: HEMATOLOGICAL MALIGANCIES
; FILE REFERENCE: 2077, 001200
; CURRENT APPLICATION NUMBER: US/09/796,692
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: 60/218,950

```

; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,378
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 9597
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (77)
; OTHER INFORMATION: n=A,T,C or G
; US-09-796-692-4917

Query Match          10.5%; Score 63; DB 9; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 AAACCCACTTACAAAACATGATGATGATCAACAATAATGTTGATGAAGAACATA 538
Db 538 AAACCCACTTACAAAACATGATGATGATGATCAACAATAATGTTGATGAAGAACATA 479

QY 599 TGA 601
Db 478 TGA 476

RESULT 43
US-10-040-862-4917/c
; Sequence 4917, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; PRIOR FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
```

```

; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (77)
; OTHER INFORMATION: n=A,T,C or G
; US-10-040-862-4917

Query Match          10.5%; Score 63; DB 14; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 AAACCCACTTACAAAACATGATGATGATCAACAATAATGTTGATGAAGAACATA 538
Db 538 AAACCCACTTACAAAACATGATGATGATGATCAACAATAATGTTGATGAAGAACATA 479

QY 599 TGA 601
Db 478 TGA 476

RESULT 44
US-10-057-475B-4917/c
; Sequence 4917, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4917
; LENGTH: 538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

/ LOCATION: (1)...(538)
/ OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-4917

Query Match 10.5%; Score 63; DB 17; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 AAACCCACTTCAAAATGATGATGATCAACAAATGTTGATGAGAGACATA 598
DB 538 AAACCCACTTCAAAATGATGATGATCAACAAATGTTGATGAGAGACATA 479

QY 599 TGA 601
DB 478 TGA 476

RESULT 45

US-10-154-884B-4917/c
/ Sequence 4917, Application US/10154884B
/ Publication No. US20040005561A1
/ GENERAL INFORMATION:

/ APPLICANT: Gaiger, Alexander
/ APPLICANT: Algate, Paul A.
/ APPLICANT: Mannion, Jane
/ APPLICANT: Retter, Marc W.
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
/ FILE REFERENCE: 014058-013521US
/ CURRENT APPLICATION NUMBER: US/10/154,884B

/ PRIOR FILING DATE: 2002-05-23
/ PRIOR APPLICATION NUMBER: US 60/186,126
/ PRIOR FILING DATE: 2000-03-01
/ PRIOR APPLICATION NUMBER: US 60/190,479
/ PRIOR FILING DATE: 2000-03-17
/ PRIOR APPLICATION NUMBER: US 60/200,545
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,303
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,779
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: US 60/200,999
/ PRIOR FILING DATE: 2000-05-01
/ PRIOR APPLICATION NUMBER: US 60/202,084
/ PRIOR FILING DATE: 2000-05-04
/ PRIOR APPLICATION NUMBER: US 60/206,201
/ PRIOR FILING DATE: 2000-05-22
/ PRIOR APPLICATION NUMBER: US 60/218,950
/ PRIOR FILING DATE: 2000-07-14
/ PRIOR APPLICATION NUMBER: US 60/222,903
/ PRIOR FILING DATE: 2000-08-03
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 11290
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 4917
/ LENGTH: 538
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(538)
/ OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-4917

Query Match 10.5%; Score 63; DB 17; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.6e-08;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 539 AAACCCACTTCAAAATGATGATGATCAACAAATGTTGATGAGAGACATA 598
DB 538 AAACCCACTTCAAAATGATGATGATCAACAAATGTTGATGAGAGACATA 479

QY 599 TGA 601
DB 478 TGA 476

Search completed: July 28, 2005, 19:13:31
Job time : 593 secs

/clone lib="NIH MGC 192"
 /note="Vector: pExpress1; Site 1: SmaI; Site 2: NotI; The library was constructed by reverse transcription of 1 ug mRNA using the oligo dt primer GCGGCGGCC(1)20 and an RNaseH + MMLV reverse transcriptase. Second strand synthesis was carried out by standard methods. The cDNA was size selected by agarose gel for > 1.2 kb, digested with Not I and directionally cloned into the vector Express-1 at the SmaI/NotI sites. DNA from the primary library was used for in vitro transcription from the T7 promoter to produce biotinylated RNA transcripts. These biotinylated transcripts, along with blocking oligos to the poly-A, multiple cloning site and primer regions, were hybridized with single stranded circles produced by phageid production from the primary library to a Cot value of 10-20. Streptavidin/phenol extraction was utilized to remove DNA:RNA hybrids leaving un-hybridized single stranded circles which were repaired by primer extension and transformed back into E. coli resulting in the normalized library. Average insert size 2.0 kb. 3 linker/adaptor sequence GCGGCGGCC(1)20. This library was constructed by Agencourt Bioscience."

ORIGIN

```

Query Match          97.9%; Score 588.4; DB 7; Length 920;
Best Local Similarity 99.7%; Pred. No. 4.1e-152;
Matches 600; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CTTGAGCGTCCGCGCGCGTCCG-CGACTCTCCGATGCGCAGATGCTTTAATTCGA 59
    |||
Db 63 CTTGAGCGTCCGCGCGCGTCCGCGCACTCTCCGATGCGCAGATGCTTTAATTCGA 122
    |||

QY 60 AAAACCTGCGCGTTCAGGACAAAGATCTTGGGTAAATGCTGCCAAATCCATCG 119
    |||
Db 123 AAAACCTGCGCGTTCAGGACAAAGATCTTGGGTAAATGCTGCCAAATCCATCG 182
    |||

QY 120 CCACCACTTAATAGACAGACAAAGTGTGCTGATGAGCTTAAAGATGACCA 179
    |||
Db 183 CCACCACTTAATAGACAGACAAAGTGTGCTGATGAGCTTAAAGATGACCA 242
    |||

QY 180 GGGAGTACACCCAAACCAAGAGGAGAGAGATCATCAAGAACCTCATCAAGACAG 239
    |||
Db 243 GGGAGTACACCCAAACCAAGAGGAGAGAGATCATCAAGAACCTCATCAAGACAG 302
    |||

QY 240 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAATCAAGATGACATGCA 299
    |||
Db 303 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAATCAAGATGACATGCA 362
    |||

QY 300 TGAAGAAATTTAAGAAAGTTTCTATGAGCTTGTATGACCGGTGATGTTCCATCAGG 359
    |||
Db 363 TGAAGAAATTTAAGAAAGTTTCTATGAGCTTGTATGACCGGTGATGTTCCATCAGG 422
    |||

QY 360 TGAATTATACCTTTGACCGGAATGTGTATTCAGAGGCTTAAATGAATGAGAGAGATGC 419
    |||
Db 423 TGAATTATACCTTTGACCGGAATGTGTATTCAGAGGCTTAAATGAATGAGAGATGC 482
    |||

QY 420 TGAACCAATCATTTGACCGGCACTCATGCGCAAGTCAATGACGGGTAAATATGTC 479
    |||
Db 483 TGAACCAATCATTTGACCGGCACTCATGCGCAAGTCAATGACGGGTAAATATGTC 542
    |||

QY 480 TTGATCATTTTTCAGATTGGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTA 539
    |||
Db 543 TTGATCATTTTTCAGATTGGAATTTTGGCTGCTGTATTAATCCTTTGGGAATTTTA 602
    |||

QY 540 AACCCCACTTACAAAACATATGATGATCAACAAAATGTTGATGAGAGAGATCAT 599
    |||
Db 603 AACCCCACTTACAAAACATATGATGATCAACAAAATGTTGATGAGAGAGATCAT 662
    |||

QY 600 GA 601
    ||
Db 663 GA 664
    ||

```

RESULT 3

LOCUS CN305329 737 bp mRNA linear EST 16-MAY-2004
 DEFINITION 1700600027114 GRN_PREHEP Homo sapiens cDNA 5', mRNA sequence.
 ACCESSION CN305329
 VERSION CN305329.1 GI:47321743
 KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 737)
 Brandenberger R., Wei H., Zhang S., Lei S., Murge J., Fisk G., J.,
 Li Y., Xu C., Pang R., Guegler K., Rao M.S., Mandalam R.,
 Lebkowski J. and Stanton L.W.

Transcriptome characterization elucidates signaling networks that
 control human ES cell growth and differentiation

Nat. Biotechnol. 22 (6) 707-716 (2004)

JOURNAL

Contact: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760

COMMENT

Email: rbrandenberger@geron.com
 Insert Length: 737 Std Error: 0.00.
 Location/Qualifiers

FEATURES

source

1..737
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="embryonic stem cells, DMSO-treated H9 cell line"
 /clone_lib="GRN PREHEP"
 /note="oligo dt primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN

```

Query Match          97.6%; Score 586.8; DB 7; Length 737;
Best Local Similarity 99.5%; Pred. No. 1.1e-151;
Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 CTTGAGCGTCCGCGCGCGTCCG-CGACTCTCCGATGCGCAGATGCTTTAATTCGA 59
    |||
Db 72 CTTGAGCGTCCGCGCGCGTCCGCGCACTCTCCGATGCGCAGATGCTTTAATTCGA 131
    |||

QY 60 AAAACCTGCGCGTTCAGGACAAAGATCTTGGGTAAATGCTGCCAAATCCATCG 119
    |||
Db 132 AAAACCTGCGCGTTCAGGACAAAGATCTTGGGTAAATGCTGCCAAATCCATCG 191
    |||

QY 120 CCACCACTTAATAGACAGACAAAGTGTGATGAGTGTGATGAGCTTCAAGAGTACCA 179
    |||
Db 192 CCACCACTTAATAGACAGACAAAGTGTGATGAGTGTGATGAGCTTCAAGAGTACCA 251
    |||

QY 180 GGGAGTACACCCAAACCAAGAGGAGAGAGATCATCAAGAACCTCATCAAGACAG 239
    |||
Db 252 GGGAGTACACCCAAACCAAGAGGAGAGAGATCATCAAGAACCTCATCAAGACAG 311
    |||

QY 240 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAATCAAGATGACATGCA 299
    |||
Db 312 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTAATCAAGATGACATGCA 371
    |||

QY 300 TGAAGAAATTTAAGAAAGTTTCTATGAGCTTGTATGACCGGTGATGTTCCATCAGG 359
    |||
Db 372 TGAAGAAATTTAAGAAAGTTTCTATGAGCTTGTATGACCGGTGATGTTCCATCAGG 431
    |||

QY 360 TGAATTATACCTTTGACCGGAATGTGTATTCAGAGGCTTAAATGAATGAGAGATGC 419
    |||
Db 432 TGAATTATACCTTTGACCGGAATGTGTATTCAGAGGCTTAAATGAATGAGAGATGC 491
    |||

QY 420 TGAACCAATCATTTGACCGGCACTCATGCGCAAGTCAATGACGGGTAAATATGTC 479
    |||

```


ORIGIN

/clone.lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

Query Match 97.6%; Score 586.8; DB 4; Length 813;
 Best Local Similarity 99.3%; Pred. No. 1.1e-151;
 Matches 599; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 CTTGAGGCTCCGGCCGCGTCCG-CGACTCTCCGAGTGGCAGAGATGCTTAATTCGA 59
 63 CTTGAGGCTCCGGCCGCGTCCGCGCACTCTCCGAGTGGCAGAGATGCTTAATTCGA 122
 60 AAAACCTGCGCGTTCAGGCAAAAAGAGATCTTGGGTAAATGCTGCCAAATCCATCG 119
 123 AAAACCTGCGCGTTCAGGCAAAAAGAGATCTTGGGTAAATGCTGCCAAATCCATCG 182
 120 CCACCACTTAAATGAGCAGACAGATGAGTGTCTGATGAGTCTTACAGAGTACCA 179
 183 CCACCACTTAAATGAGCAGACAGATGAGTGTCTGATGAGTCTTACAGAGTACCA 242
 180 GGGAGTACACCCAAACAAAGAGGAGAGAGATCATCAAGAACCTCATCAACAG 239
 243 GGGAGTACACCCAAACAAAGAGGAGAGAGATCATCAAGAACCTCATCAACAG 302
 240 TCATCAAGCTGGCATTCTTTATAGCAATATCAGTTTATCAAGATGAGTCAATGGA 299
 303 TCATCAAGCTGGCATTCTTTATAGCAATATCAGTTTATCAAGATGAGTCAATGGA 362
 300 TGGAGAAATTTAAGAAAGATTCATCAGCTTCTATGACCGTGTGATGTTCCATCAGG 359
 363 TGGAGAAATTTAAGAAAGATTCATCAGCTTCTATGACCGTGTGATGTTCCATCAGG 422
 360 TGGATTAATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAAATGAGAGATGC 419
 423 TGGATTAATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAAATGAGAGATGC 482
 420 TGGACCAAAATCATTTAGCGGCACTCATCTGCAAGTCAATGAGCGGTTAATATGTCT 479
 483 TGGACCAAAATCATTTAGCGGCACTCATCTGCAAGTCAATGAGCGGTTAATATGTCT 542
 480 TTGATCAATTTTTCAGATTGGAATTTTGGCTGCTGTATTAATCTTTTGGGAATTTTA 539
 543 TTGATCAATTTTTCAGATTGGAATTTTGGCTGCTGTATTAATCTTTTGGGAATTTTA 602
 540 AACCCCACTTACAAAACCTATGTGATGTGTATCAACAAAATGTTGATGAAGAGATCAT 599
 603 AACCCCACTTACAAAACCTATGTGATGTGTATCAACAAAATGTTGATGAAGAGATCAT 662
 600 GA 601
 663 GA 664

RESULT 6
 BO424670
 LOCUS BO424670
 DEFINITION AGENCOURT_7833747 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6153217
 5', mRNA sequence.
 ACCESSION BO424670
 VERSION BO424670.1 GI:21119985
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 1 (bases 1 to 896)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM13492 row: F column: 02
 High quality sequence stop: 596.

FEATURES

source
 1..896
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6153217"
 /issue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_67"
 /note="Organ: eye; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 kb. Library constructed by Life Technologies."

ORIGIN

Query Match 97.5%; Score 585.8; DB 5; Length 896;
 Best Local Similarity 99.3%; Pred. No. 2.1e-151;
 Matches 598; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

1 CTTGAGGCTCCGGCCGCGTCCG-CGACTCTCCGAGTGGCAGAGATGCTTAATTCGA 59
 62 CTTGAGGCTCCGGCCGCGTCCGCGCACTCTCCGAGTGGCAGAGATGCTTAATTCGA 121
 60 AAAACCTGCGCGTTCAGGCAAAAAGAGATCTTGGGTAAATGCTGCCAAATCCATCG 119
 122 AAAACCTGCGCGTTCAGGCAAAAAGAGATCTTGGGTAAATGCTGCCAAATCCATCG 181
 120 CCACCACTTAAATGAGCAGACAGATGAGTGTCTGATGAGTCTTACAGAGTACCA 179
 182 CCACCACTTAAATGAGCAGACAGATGAGTGTCTGATGAGTCTTACAGAGTACCA 241
 180 GGGAGTACACCCAAACAAAGAGGAGAGAGATCATCAAGAACCTCATCAACAG 239
 242 GGGAGTACACCCAAACAAAGAGGAGAGAGATCATCAAGAACCTCATCAACAG 301
 240 TCATCAAGCTGGCATTCTTTATAGCAATATCAGTTTATCAAGATGAGTCAATGGA 299
 302 TCATCAAGCTGGCATTCTTTATAGCAATATCAGTTTATCAAGATGAGTCAATGGA 361
 300 TGGAGAAATTTAAGAAAGATTCATCAGCTTCTATGACCGTGTGATGTTCCATCAGG 359
 362 TGGAGAAATTTAAGAAAGATTCATCAGCTTCTATGACCGTGTGATGTTCCATCAGG 421
 360 TGGATTAATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAAATGAGAGATGC 419
 422 TGGATTAATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAAATGAGAGATGC 481
 420 TGGACCAAAATCATTTAGCGGCACTCATCTGCAAGTCAATGAGCGGTTAATATGTCT 479
 482 TGGACCAAAATCATTTAGCGGCACTCATCTGCAAGTCAATGAGCGGTTAATATGTCT 541
 480 TTGATCAATTTTTCAGATTGGAATTTTGGCTGCTGTATTAATCTTTTGGGAATTTTA 539
 542 TTGATCAATTTTTCAGATTGGAATTTTGGCTGCTGTATTAATCTTTTGGGAATTTTA 601
 540 AACCCCACTTACAAAACCTATGTGATGTGTATCAACAAAATGTTGATGAAGAGATCAT 599

Db 602 AACCCACTTACAAAAATATGATGATGATCAACAAATGTTGATGAAGAGACATAT 661
QY 600 GA 601
Db 662 GA 663

RESULT 7
BI752550 876 bp mRNA linear EST 25-SEP-2001
LOCUS 603021969F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5192561 5',
DEFINITION mRNA sequence.
ACCESSION BI752550
VERSION BI752550.1 GI:15744128
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 876)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-rcmail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINT at:
http://image.llnl.gov
Plate: LLM11481 row: 3 column: 18
High quality sequence stop: 813.
Location/Qualifiers
1. 876
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5192561"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_114"
/note="Organ: brain; Vector: PCMV-SPORE6; Site 1: NotI;
Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 019. Note:
this is a NIH_MGC Library."

ORIGIN

Query Match 95.5%; Score 574.2; DB 4; Length 876;
Best Local Similarity 99.2%; Pred. No. 3.5e-148;
Matches 558; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 1 CTTGACGCTCCGCGCCGCTGCG-CACTCTCCGATGCGCAGACAGATGCTTTAATCCA 59
Db 83 CTTGACGCTCCGCGCCGCTGCGCAGCTCTCCGATGCGCAGATGCTTTAATCCA 142

QY 60 AAAAAGCTGCGCTTCAAGGCAAAAGAAAGTCTGGGTAAATGCTGCCAATTCATCG 119
Db 143 AAAAAGCTGCGCTTCAAGGCAAAAGAAAGTCTGGGTAAATGCTGCCAATTCATCG 202

QY 120 CCAACCACTTAATAGACGACACAAGTAGTGTGCTGATGAGCTTACAGAGTACCA 179
Db 203 CCAACCACTTAATAGACGACACAAGTAGTGTGCTGATGAGCTTACAGAGTACCA 262

QY 180 GGGAGTACACCCAAACAAAGAGAGGCGAGAGATCATCAAGAACTTCATCAAGACAG 239
Db 263 GGGAGTACACCCAAACAAAGAGAGGCGAGAGATCATCAAGAACTTCATCAAGACAG 322

QY 240 TCATCAGCTGCGCATTCCTTTATAGAAATATCAGTTAATCAAGATGAGTACATTGA 299
Db 323 TCATCAGCTGCGCATTCCTTTATAGAAATATCAGTTAATCAAGATGAGTACATTGA 382

QY 300 TGGAGAAATTTAAGAAAGATTTCATCAGCTTGTATGACCGTGTGCTGATTCATCAG 359
Db 383 TGGAGAAATTTAAGAAAGATTTCATCAGCTTGTATGACCGTGTGCTGATTCATCAG 442

QY 360 TGGATTATACCTTTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGCAGAGATGC 419
Db 443 TGGATTATACCTTTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGCAGAGATGC 502

QY 420 TGCACCAATCATTTGACGCGCCACCTCAGTCCAAATCATAGAGCGGTTAATATGCT 479
Db 503 TGCACCAATCATTTGACGCGCCACCTCAGTCCAAATCATAGAGCGGTTAATATGCT 562

QY 480 TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATCC-TTTTGGGAATTT 538
Db 563 TTGATCATTTTTCAGATTGTGAATTTTGGCTGCTGTATATCC-TTTTGGGAATTT 622

QY 539 AAACCCCACTTACAAAACATATGTGATGATCAACAAAATTTGATGACAGAACATA 598
Db 623 AAACCCCACTTACAAAACATATGTGATGATCAACAAAATTTGATGACAGAACATA 682

QY 599 TGA 601
Db 683 TGA 685

RESULT 8
CN305327 712 bp mRNA linear EST 16-MAY-2004
LOCUS 17000532623591 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
DEFINITION CN305327
ACCESSION CN305327
VERSION CN305327.1 GI:47321741
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 712)
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,
Li, Y., Xu, C., Fang, R., Guejler, K., Rao, M.S., Mandalam, R.,
Lebowksi, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 712 Std Error: 0.00.
Location/Qualifiers
1. 712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cbase_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_1ib="GRN_ES"
/note="oligo dT primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Query Match 93.9%; Score 564.4; DB 7; Length 712;
Best Local Similarity 99.8%; Pred. No. 1.8e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCAAAACCTTGCGCGCTTCAGGACAAAAGAGATCTTGG 95

```

Db      82  TGGCCACAGATGCTTTAAATTCAGAAAACCTGGCCGCTTCAGGACAAAAGAGATCTTGG 141
Qy      96  GTTAAATGCTGTCCAAATTCATGCGCCACCACTTAAATAGACACAAAGTAGAGTGC 155
Db      142  GTTAAATGCTGTCCAAATTCATGCGCCACCACTTAAATAGACACAAAGTAGAGTGC 201
Qy      156  TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCCAAAACAAGAGAGGACAGAGAGA 215
Db      202  TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCCAAAACAAGAGAGGACAGAGAGA 261
Qy      216  TCATCAAGAACTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCAAT 275
Db      262  TCATCAAGAACTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCAAT 321
Qy      276  TTAATCAAGATGAGTACATGATGAGAGAAATTTAAGAAAGTTCACTGCTGCTA 335
Db      322  TTAATCAAGATGAGTACATGATGAGAGAAATTTAAGAAAGTTCACTGCTGCTA 381
Qy      336  TGAACCGTGTCACTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTGTATCCAGGC 395
Db      382  TGAACCGTGTCACTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTGTATCCAGGC 441
Qy      396  TGTAAATGATGATGACAGAGATGCTGACCAAAATCATTGACGCGCACTGACGCAAGT 455
Db      442  TGTAAATGATGATGACAGAGATGCTGACCAAAATCATTGACGCGCACTGACGCAAGT 501
Qy      456  CACATGAGCGGGTTAATATATGCTTGTGATCATTTTTCAGATTGTAATTTTGGCTGCT 515
Db      502  CACATGAGCGGGTTAATATATGCTTGTGATCATTTTTCAGATTGTAATTTTGGCTGCT 561
Qy      516  TGTATATCTCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 575
Db      562  TGTATATCTCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 621
Qy      576  AAATGTTGATGAGAGAAACATATGA 601
Db      622  AAATGTTGATGAGAGAAACATATGA 647

```

```

RESULT 9 718 bp  mRNA  linear  EST 25-MAR-2004
AL549492  LOCUS
DEFINITION  AL549492 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
ACCESSION  AL549492
VERSION  AL549492.3 GI:45749887
KEYWORDS  EST.
SOURCE  Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 718)
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31271310.
COMMENT  Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
3485.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CSOD1044AD070P1ec=3485.f.
Location/Qualifiers
1..718
/organism="Homo sapiens"
/mol_type="mRNA"

```

```

/db_xref="taxon:9606"
/clone="CSOD1044YG13"
/issue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

```

Query Match 93.9%; Score 564.4; DB 1; Length 718;
Best Local Similarity 99.8%; Pred. No. 1,8e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      36  TGGCCACAGATGCTTTAAATTCAGAAAACCTGGCCGCTTCAGGACAAAAGAGATCTTGG 95
Db      1  TGGCCACAGATGCTTTAAATTCAGAAAACCTGGCCGCTTCAGGACAAAAGAGATCTTGG 60
Qy      96  GTTAAATGCTGTCCAAATTCATGCGCCACCACTTAAATAGACACAAAGTAGAGTGC 155
Db      61  GTTAAATGCTGTCCAAATTCATGCGCCACCACTTAAATAGACACAAAGTAGAGTGC 120
Qy      156  TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCCAAAACAAGAGAGGACAGAGAGA 215
Db      121  TGAATGAGCTCTACAGAGTACCAAGGAGTACACCCCAAAACAAGAGAGGACAGAGAGA 180
Qy      216  TCATCAAGAACTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCAAT 275
Db      181  TCATCAAGAACTCATCAAGACAGTCAAGCTGGCCATCTTTATAGGAATATCAAT 240
Qy      276  TTAATCAAGATGAGTACATGATGAGAGAAATTTAAGAAAGTTCACTGCTGCTA 335
Db      241  TTAATCAAGATGAGTACATGATGAGAGAAATTTAAGAAAGTTCACTGCTGCTA 300
Qy      336  TGAACCGTGTCACTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTGTATCCAGGC 395
Db      301  TGAACCGTGTCACTTTCCATCAGGTGATTAATACCTTTGACCGGAATGTGTATCCAGGC 360
Qy      396  TGTAAATGATGATGACAGAGATGCTGACCAAAATCATTGACGCGCACTGACGCAAGT 455
Db      361  TGTAAATGATGATGACAGAGATGCTGACCAAAATCATTGACGCGCACTGACGCAAGT 420
Qy      456  CACATGAGCGGGTTAATATATGCTTGTGATCATTTTTCAGATTGTAATTTTGGCTGCT 515
Db      421  CACATGAGCGGGTTAATATATGCTTGTGATCATTTTTCAGATTGTAATTTTGGCTGCT 480
Qy      516  TGTATATCTCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 575
Db      481  TGTATATCTCTTTTGGGAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 540
Qy      576  AAATGTTGATGAGAGAAACATATGA 601
Db      541  AAATGTTGATGAGAGAAACATATGA 566

```

```

RESULT 10 731 bp  mRNA  linear  EST 06-JUN-2003
CD521721  LOCUS
DEFINITION  CD521721 AGNCOURT 14353034 NIH-MGC_191 Homo sapiens cDNA clone
ACCESSION  CD521721
VERSION  CD521721.1 GI:31453439
KEYWORDS  EST.
SOURCE  Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 731)
NIH-MGC http://mgi.cni.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics

```

National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Narayan Bhat
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: NDCM207 row: e column: 11
 High quality sequence stop: 603.
 Location/Qualifiers
 1. 731
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30413002"
 /issue_type="Pooled"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH MGC 191"
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccatcagcc);
 Site 2: SfiI (ggcgccctcgcc); Library is oligo-dT primed
 and directionally cloned. PBMC - Peripheral Blood
 Mononuclear Cells. RNA was pooled from 3/6hour stimulation
 with PMA adn Ionomycin. 5' and 3' adaptors were used in
 cloning as follows: 5' adaptor sequence:
 5'-ATCTGAGGCGGAGCGGCCGACAG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.69
 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 93.9%; Score 564.4; DB 6; Length 731;
 Best Local Similarity 99.8%; Pred. No. 1.8e-145;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGCAGCAAAAGAGATCTTG 95
 31 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGCAGCAAAAGAGATCTTG 90
 96 GTPAATGAGTCTTCAATCATCGCCACCACTTAATGAGCAGACAGTATGAGTGC 155
 91 GTPAATGAGTCTTCAATCATCGCCACCACTTAATGAGCAGACAGTATGAGTGC 150
 156 TGGATGAGTCTTCAAGAGTACAGAGGAGTACACCAAAACAAGAGAGAGAGAGA 215
 151 TGGATGAGTCTTCAAGAGTACAGAGGAGTACACCAAAACAAGAGAGAGAGAGA 210
 216 TCATCAAGAACTCATCAAGAGTCAAGCTGGCCATTTTATAGGAATATCAAGT 275
 211 TCATCAAGAACTCATCAAGAGTCAAGCTGGCCATTTTATAGGAATATCAAGT 270
 276 TTAATCAAGTGAAGTGAAGTGAAGAAATTTAAGAAAGTTCATCACTGCTGTA 335
 271 TTAATCAAGTGAAGTGAAGTGAAGAAATTTAAGAAAGTTCATCACTGCTGTA 330
 336 TGAACGCTGTCAGTTCATCAAGTGAATATCACTTGAACCGGAATGTTTATCAAGC 395
 331 TGAACGCTGTCAGTTCATCAAGTGAATATCACTTGAACCGGAATGTTTATCAAGC 390
 396 TGTAAATGATGACAGAGATCTGACCAATCATTCAGCGCCACTCATGCAAGT 455
 391 TGTAAATGATGACAGAGATCTGACCAATCATTCAGCGCCACTCATGCAAGT 450
 456 CACATGAGAGGTTAATTAATGCTTGAATTTTCAATGTTGAAATTTTGGCTGCT 515
 451 CACATGAGAGGTTAATTAATGCTTGAATTTTCAATGTTGAAATTTTGGCTGCT 510
 516 TGTATATCTTTTGGAAATTTAAACCCCACTTACAAAAAATATGATGATATCAACA 575

Db 511 TGTATATCTTTTGGAAATTTTAAACCCCACTTACAAAAAATATGATGATATCAACA 570
 Qy 576 AATGTTGATGAAGAGAAATATGA 601
 Db 571 AATGTTGATGAAGAGAAATATGA 596

RESULT 11

AL527566 897 bp mRNA linear EST 24-MAR-2004
 LOCUS AL527566 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DC024YA23 5-PRIME, mRNA sequence.
 ACCESSION AL527566
 VERSION AL527566.3 GI:45702664
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 897)
 Li W.B., Gruber,C., Jessee,J. and Polayes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On Feb 13, 2001 this sequence version replaced gi:31065417.
 COMMENT
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 3485.f

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0DC024YA12QPI&c=3485.f.

FEATURES

source

1. 897
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DC024YA23"
 /issue_type="NEUROBLASTOMA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 93.9%; Score 564.4; DB 1; Length 897;
 Best Local Similarity 99.8%; Pred. No. 1.9e-145;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGCAGCAAAAGAGATCTTG 95
 1 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCCAGCAGCAAAAGAGATCTTG 60
 96 GTPAATGAGTCTTCAATCATCGCCACCACTTAATGAGCAGACCAATGATGAGTGC 155
 61 GTPAATGAGTCTTCAATCATCGCCACCACTTAATGAGCAGACCAATGATGAGTGC 120
 156 TGGATGAGTCTTCAAGAGTACAGAGGAGTACACCAAAACAAGAGAGAGAGAGA 215
 121 TGGATGAGTCTTCAAGAGTACAGAGGAGTACACCAAAACAAGAGAGAGAGAGA 180
 216 TCATCAAGAACTCATCAAGAGTCAAGCTGGCCATTTTATAGGAATATCAAGT 275
 181 TCATCAAGAACTCATCAAGAGTCAAGCTGGCCATTTTATAGGAATATCAAGT 240
 276 TTAATCAAGTGAAGTGAAGTGAAGAAATTTAAGAAAGTTCATCACTGCTGTA 335

Db 241 TTAATCAAGATGAGTACGATGATGAGAAATTTAAGAAAGATTCATCAGCTGCTA 300
 QY 336 TGACCGTGGTCACTTTCATCAGTGTATTAACCTTTGACCGGAATGTGTATTCAGGC 395
 Db 301 TGACCGTGGTCACTTTCATCAGTGTATTAACCTTTGACCGGAATGTGTATTCAGGC 360
 QY 396 TGTAAATGAATCAGAGATGCTGCACCAATTCATTGAGCGGCACCTCAGTCCAGT 455
 Db 361 TGTAAATGAATCAGAGATGCTGCACCAATTCATTGAGCGGCACCTCAGTCCAGT 420
 QY 456 CACATGACGGGTTAAATATGTCTTTCATCATTTCAGATTGTGAATTTTGGCTGCT 515
 Db 421 CACATGACGGGTTAAATATGTCTTTCATCATTTCAGATTGTGAATTTTGGCTGCT 480
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGTATTCACA 575
 Db 481 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGTATTCACA 540
 QY 576 AAATGTTGATGAAGAAACATATGA 601
 Db 541 AAATGTTGATGAAGAAACATATGA 566

RESULT 12
 BM919223 1055 bp mRNA linear EST 12-MAR-2002
 LOCUS BM919223
 DEFINITION AGENCOURT_6715690 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748442
 5', mRNA sequence.

ACCESSION BM919223
 VERSION BM919223.1 GI:19369602
 KEYWORDS EST
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Makiyoda; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL NIH-MGC http://mgi.nci.nih.gov/
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM12776 row: d column: 11
 High quality sequence stop: 677.
 Location/Qualifiers

FEATURES
 source
 1..1055
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5748442"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_120"
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 93.8%; Score 564.4; DB 5; Length 1055;
 Best Local Similarity 99.8%; Pred. No. 2e-145;
 Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGCCACAGATGCTTAAATTCAAAACCTGCGCTTACAGCAGCAAAAAGATCTTG 95
 Db 33 TGCCACAGATGCTTAAATTCAAAACCTGCGCTTACAGCAGCAAAAAGATCTTG 92
 QY 96 GTAAATGCTGTCAAATTCATGCGCACCACCTTAATAGACACAAAGATGAGGTG 155
 Db 93 GTAAATGCTGTCAAATTCATGCGCACCACCTTAATAGACACAAAGATGAGGTG 152
 QY 156 TGGATGAGCTTACAGAGTACCGAGAGTACACCCAAAACAAAGAGAGCAGAGA 215
 Db 153 TGGATGAGCTTACAGAGTACCGAGAGTACACCCAAAACAAAGAGAGCAGAGA 212
 QY 216 TCAATGAAGAACTTCATCAAGACAGTCAAGTGGCCATTCTTTATAGAAATACAG 275
 Db 213 TCAATGAAGAACTTCATCAAGACAGTCAAGTGGCCATTCTTTATAGAAATACAG 272
 QY 276 TTAATCAAGATGAGTATGATGATGAGAAATTTAAGAAAGATTCACAGTCTGCTA 335
 Db 273 TTAATCAAGATGAGTATGATGATGAGAAATTTAAGAAAGATTCACAGTCTGCTA 332
 QY 336 TGACCGTGGTCACTTTCATCAGTGTATTAACCTTTGACCGGAATGTGTATTCAGGC 395
 Db 333 TGACCGTGGTCACTTTCATCAGTGTATTAACCTTTGACCGGAATGTGTATTCAGGC 392
 QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTAGCGGCCCTCAGTCCAGT 455
 Db 393 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTAGCGGCCCTCAGTCCAGT 452
 QY 456 CACATGACGGGTTAAATATGTCTTTCATCATTTCAGATTGTGAATTTTGGCTGCT 515
 Db 453 CACATGACGGGTTAAATATGTCTTTCATCATTTCAGATTGTGAATTTTGGCTGCT 512
 QY 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGTATTCACA 575
 Db 513 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGTATTCACA 572
 QY 576 AAATGTTGATGAAGAAACATATGA 601
 Db 573 AAATGTTGATGAAGAAACATATGA 598

RESULT 13
 CR611795 1610 bp mRNA linear HTC 21-JUL-2004
 LOCUS CR611795
 DEFINITION full-length cDNA clone CS01057YA07 of Placenta Cot 25-normalized
 of Homo sapiens (human).
 ACCESSION CR611795
 VERSION CR611795.1 GI:50492602
 KEYWORDS HTC; CNSLT_cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Makiyoda; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 JOURNAL NIH-MGC http://mgi.nci.nih.gov/
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Feng Liang Email: fliang@life.rockefeller.edu
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue
 Gaithersburg, MD 20878
 2 (bases 1 to 1610)
 Genoscope.

REFERENCE
 AUTHORS Direct Submission
 TITLE Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

FEATURES
 source
 Location/Qualifiers
 1..1610

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS001057YA07"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 93.9%; Score 564.4; DB 3; Length 1610;
Best Local Similarity 99.8%; Pred. No. 2.2e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTGAGGACAGAAAGAGATCTTG 95
DB 1 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTGAGGACAGAAAGAGATCTTG 60
QY 96 GTAAATGGTGTCCAAATCCATGCGCAACCTTAATGACGACACAGATAGTGGTC 155
DB 61 GTAAATGGTGTCCAAATCCATGCGCAACCTTAATGACGACACAGATAGTGGTC 120
QY 156 TGGATGAGCTCTACAGAGTGACCGAGGAGTACCCCAAAACAGAGAGGAGAGAGA 215
DB 121 TGGATGAGCTCTACAGAGTGACCGAGGAGTACCCCAAAACAGAGAGGAGAGAGA 180
QY 216 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTAATAGGAATATCACT 275
DB 181 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTAATAGGAATATCACT 240
QY 276 TTAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTCAATCAGCTTGCTA 335
DB 241 TTAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTCAATCAGCTTGCTA 300
QY 336 TGACCGTGTGCTTCCATCAGTGAGTATATACCTTTGACCGGAATGTTATCCAGGC 395
DB 301 TGACCGTGTGCTTCCATCAGTGAGTATATACCTTTGACCGGAATGTTATCCAGGC 360
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGCCACTCATCTGCCAAGT 455
DB 361 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGCCACTCATCTGCCAAGT 420
QY 456 CACATGAGAGGGTTAATAATGCTTTGATCATTTTCAAGTTGTGAATTTTGGCTGCT 515
DB 421 CACATGAGAGGGTTAATAATGCTTTGATCATTTTCAAGTTGTGAATTTTGGCTGCT 480
QY 516 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGATCAACA 575
DB 481 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGATCAACA 540
QY 576 AAATGTTGATGAAGAACAATATGA 601
DB 541 AAATGTTGATGAAGAACAATATGA 566

RESULT 14
CR615621 1733 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS00C024YA23 of Neuroblastoma Cot
DEFINITION 25-normalized of Homo sapiens (human).
ACCESSION CR615621
VERSION CR615621.1 GI:50496428
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact : Feng Liang Email : fliang@life.uchicago.edu
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
REFERENCE 2 (bases 1 to 1733)

AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)

COMMENT

1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source

Location/Qualifiers
1..1733

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS00C024YA23"
/issue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 93.9%; Score 564.4; DB 3; Length 1733;
Best Local Similarity 99.8%; Pred. No. 2.2e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTGAGGACAGAAAGAGATCTTG 95
DB 1 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTGAGGACAGAAAGAGATCTTG 60
QY 96 GTAAATGGTGTCCAAATCCATGCGCAACCTTAATGACGACACAGATAGTGGTC 155
DB 61 GTAAATGGTGTCCAAATCCATGCGCAACCTTAATGACGACACAGATAGTGGTC 120
QY 156 TGGATGAGCTCTACAGAGTGACCGAGGAGTACCCCAAAACAGAGAGGAGAGAGA 215
DB 121 TGGATGAGCTCTACAGAGTGACCGAGGAGTACCCCAAAACAGAGAGGAGAGAGA 180
QY 216 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTAATAGGAATATCACT 275
DB 181 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTAATAGGAATATCACT 240
QY 276 TTAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTCAATCAGCTTGCTA 335
DB 241 TTAATCAAGATGAGCTAGCATGATGAGAAATTTAAGAAAGATTCAATCAGCTTGCTA 300
QY 336 TGACCGTGTGCTTCCATCAGTGAGTATATACCTTTGACCGGAATGTTATCCAGGC 395
DB 301 TGACCGTGTGCTTCCATCAGTGAGTATATACCTTTGACCGGAATGTTATCCAGGC 360
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGCCACTCATCTGCCAAGT 455
DB 361 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGCCACTCATCTGCCAAGT 420
QY 456 CACATGAGAGGGTTAATAATGCTTTGATCATTTTCAAGTTGTGAATTTTGGCTGCT 515
DB 421 CACATGAGAGGGTTAATAATGCTTTGATCATTTTCAAGTTGTGAATTTTGGCTGCT 480
QY 516 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGATCAACA 575
DB 481 TGTAAATGCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGATCAACA 540
QY 576 AAATGTTGATGAAGAACAATATGA 601
DB 541 AAATGTTGATGAAGAACAATATGA 566

RESULT 15
CR615131 1754 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS001081Y011 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR615131
VERSION CR615131.1 GI:50495938
KEYWORDS HTC; CNSLT_cDNA.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 1754)
TITLE L1.W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
REMARK Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue
Genoscope.
2 (bases 1 to 1754)
REFERENCE Direct Submission
AUTHORS Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
JOURNAL - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
source
1..1754
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1081Y011"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

Query Match 93.9%; Score 564.4; DB 3; Length 1754;
Best Local Similarity 99.8%; Pred. No. 2.2e-145;
Matches 565; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAAATTCGCAAAACCTGGCCGTTGCGCAAAAGAGATCTTGG 95
DB 1 TGGCCACAGATGCTTTAAATTCGCAAAACCTGGCCGTTGCGCAAAAGAGATCTTGG 60

QY 96 GTAAATGATGCTCAATTCATCGCCACCACTTAATAGACACAAAGTAGAGGTG 155
DB 61 GTAAATGATGCTCAATTCATCGCCACCACTTAATAGACACAAAGTAGAGGTG 120

QY 156 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCAAGAGAGAGGAGAGAGA 215
DB 121 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCAAGAGAGAGGAGAGAGA 180

QY 216 TCATCAAGAACCTCATCAAGACAGTCAAGCTGCGCATTTTATAGAAATATCAGT 275
DB 181 TCATCAAGAACCTCATCAAGACAGTCAAGCTGCGCATTTTATAGAAATATCAGT 240

QY 276 TTAATCAAGATGAGCTGAGTATGATGAGAAATTAAGAAAGATTCACAGCTGCTA 335
DB 241 TTAATCAAGATGAGCTGAGTATGATGAGAAATTAAGAAAGATTCACAGCTGCTA 300

QY 336 TGAACCGTGTCAAGTTCATCAGGTGATTAACCTTGAACCGGATGTGTTATCCAGGC 395
DB 301 TGAACCGTGTCAAGTTCATCAGGTGATTAACCTTGAACCGGATGTGTTATCCAGGC 360

QY 396 TGTAAATGATGACAGAGATGCTGACCAAAATCATTCAGCGCACCTCAGTCCAGT 455
DB 361 TGTAAATGATGACAGAGATGCTGACCAAAATCATTCAGCGCACCTCAGTCCAGT 420

QY 456 CACATGAGCGGTTAATATGCTTTGATCATTTTTCAGATTTGGAATTTTGGCTGCT 515
DB 421 CACATGAGCGGTTAATATGCTTTGATCATTTTTCAGATTTGGAATTTTGGCTGCT 480

QY 516 TGTATAATCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGATGATGATCAACA 575
DB 481 TGTATAATCTTTTGGGAATTTTAAACCCCACTTACAAAAAATATGATGATGATGATCAACA 540

QY 576 AAATGTTGATGAGAGACATATGA 601

DB 541 AAATGTTGATGAGAGACATATGA 566

RESULT 16
AL550457
LOCUS
DEFINITION
AL550457 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1057YA07 5-PRIME, mRNA sequence.
ACCESSION
AL550457
VERSION
AL550457.3 GI:45750828
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS 1 (bases 1 to 885)
JOURNAL L1.W.B., Gruber,C., Jessee,J. and Polayes,D.
REMARK Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31272274.
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3485.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0D1057YA040P1c=3485.f.
FEATURES
source
1..885
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1057YA07"
/issue_type="Placenta COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 93.6%; Score 562.8; DB 1; Length 885;
Best Local Similarity 99.6%; Pred. No. 5.2e-145;
Matches 564; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAAATTCGCAAAACCTGGCCGTTGCGCAAAAGAGATCTTGG 95
DB 1 TGGCCACAGATGCTTTAAATTCGCAAAACCTGGCCGTTGCGCAAAAGAGATCTTGG 60

QY 96 GTAAATGATGCTCAATTCATCGCCACCACTTAATAGACACAAAGTAGAGGTG 155
DB 61 GTAAATGATGCTCAATTCATCGCCACCACTTAATAGACACAAAGTAGAGGTG 120

QY 156 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCAAGAGAGAGGAGAGAGA 215
DB 121 TGGATGAGCTCTACAGAGTACCAAGGAGTACACCAAGAGAGAGGAGAGAGA 180

QY 216 TCATCAAGAACCTCATCAAGACAGTCAAGCTGCGCATTTTATAGAAATATCAGT 275
DB 181 TCATCAAGAACCTCATCAAGACAGTCAAGCTGCGCATTTTATAGAAATATCAGT 240

QY 276 TTAATCAAGATGAGCTGAGTATGATGAGAAATTAAGAAAGATTCACAGCTGCTA 335
DB 241 TTAATCAAGATGAGCTGAGTATGATGAGAAATTAAGAAAGATTCACAGCTGCTA 300

QY 336 TGAACCGTGTCAAGTTCATCAGGTGATTAACCTTGAACCGGATGTGTTATCCAGGC 395

Db 301. TGAACGTCGTCAGTTTCATCAGTGTGATTAATACCTTTGACCGGAAATGTTATCCAGGC 360

Qy 396 TGTAAATGATGACAGAGATGCTGACCAATTCATTGAGGCGCACTCAGTCCAGT 455

Db 361 TGTAAATGATGACAGAGATGCTGACCAATTCATTGAGGCGCACTCAGTCCAGT 420

Qy 456 CACATGACGGGTAAATGATCTTTGATCATTTCAGATTGTGAATTTTGGCTGCT 515

Db 421 CACATGACGGGTAAATGATCTTTGATCATTTCAGATTGTGAATTTTGGCTGCT 480

Qy 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTATGATCAACA 575

Db 481 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTATGATCAACA 540

Qy 576 AAATGTTGATGAGAGAACATATGA 601

Db 541 AAATGTTGATGAGAGAACATATGA 566

RESULT 17

AL554168 1045 bp mRNA linear EST 30-MAR-2004

LOCUS AL554168 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

DEFINITION clone CS0D1081Y011 5-PRIME, mRNA sequence.

ACCESSION AL554168

VERSION AL554168.3 GI:45858929

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1045)

AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On Feb 15, 2001 this sequence version replaced gi:31275981.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 3485.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?c=CS0D1081AH06QPlc=3485.f.

FEATURES

source

1..1045

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0D1081Y011"

/tissue_type="PLACENTA COT 25-NORMALIZED"

/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 93.6%; Score 562.4; DB 1; Length 1045;

Best Local Similarity 99.5%; Pred. No. 7e-145; 2; Indels 0; Gaps 0;

Matches 563; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 36 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTCACGACAAAAGAGATCTTG 95

Db 1 TGGCCACAGATGCTTTAATTCAAAACCTGGCCGTCACGACAAAAGAGATCTTG 60

Qy 96 GTRAAATGATGTCACAAATCATTCGACCAACCTTATATAGACACAAAGTAGAGAGTGC 155

Db 61 GTRAAATGATGTCACAAATCATTCGACCAACCTTATATAGACACAAAGTAGAGTGC 120

Qy 156 TGTATAGCTCTACAGAGTACAGAGAGTACACCCAAAACAGAGAGGACAGAGA 215

Db 121 TGTATAGCTCTACAGAGTACAGAGAGTACACCCAAAACAGAGAGGACAGAGA 180

Qy 216 TCATCAAGAACCTCATCAAGACATCATCAAGCTGCGCATTTTATAGAAATATCAGT 275

Db 181 TCATCAAGAACCTCATCAAGACATCATCAAGCTGCGCATTTTATAGAAATATCAGT 240

Qy 276 TTAATCAAGAACCTCATCAAGTATGAGAGAAATTTAAGAAAGAAATGATCAGCTGCTA 335

Db 241 TTAATCAAGAACCTCATCAAGTATGAGAGAAATTTAAGAAAGAAATGATCAGCTGCTA 300

Qy 336 TGACCGTGTGTCAGTTTCATCAGTGTGATTTATACCTTTGACCGGATGTATCAGGC 395

Db 301 TGACCGTGTGTCAGTTTCATCAGTGTGATTTATACCTTTGACCGGATGTATCAGGC 360

Qy 396 TGTAAATGATGACAGAGATGCTGACCAAAATCATTCAGGCGCACTCAGTCCAGT 455

Db 361 TGTAAATGATGACAGAGATGCTGACCAAAATCATTCAGGCGCACTCAGTCCAGT 420

Qy 456 CACATGACGGGTAAATGATCTTTGATCATTTCAGATTGTGAATTTTGGCTGCT 515

Db 421 CACATGACGGGTAAATGATCTTTGATCATTTCAGATTGTGAATTTTGGCTGCT 480

Qy 516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTATGATCAACA 575

Db 481 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGTATGATCAACA 540

Qy 576 AAATGTTGATGAGAGAACATATGA 601

Db 541 AAATGTTGATGAGAGAACATATGA 566

RESULT 18

CB991862 785 bp mRNA linear EST 01-MAY-2003

LOCUS CB991862

DEFINITION AGENCOURT 13618130 NIH MGC 148 Homo sapiens cDNA clone

IMAGE:30337918 5', mRNA sequence.

ACCESSION CB991862

VERSION CB991862.1 GI:30286382

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 785)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Dr. Stefan Hanson

cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help and advice from Piero Carninci (RIKEN)

cDNA library Arrayed by: The I.M.A.G.E. Consortium (ILNI)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at:

http://image.llnl.gov

Plate: NDM363 row: 1 column: 23

High quality sequence stop: 629.

Location/Qualifiers

1..785

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30337918"

/tissue_type="pre-eclampic placenta"

/lab_host="DH10B Tona"

/clone_lib="NIH_MGC_148"

/note="Organ: Placenta; Vector: pBluescriptR, Site_1: all-XhoI; Site_2: BamH; Library: is oligo-dT primed and

directionally cloned using primer
5'-TTTTTTTTTTTTTTT-3', size-selected for average insert
size 2.3 kb and normalized to 10⁷ 5'. This is a primary
library enriched for full-length clones and constructed
using the Cap-trapper method (Carninci, in preparation).
Library constructed by M. Brownstein (NIMH/NHRI,
National Institutes of Health). Note: this is a NIH-MGC
Library."

ORIGIN

Query Match 93.2%; Score 560.4; DB 6; Length 785;
Best Local Similarity 99.5%; Pred. No. 2,36-144;
Matches 583; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 CTTGAGGTCCTGGGCGGCGGTCGCGGCTCTCCGATGCGCAGATGCTTTAATTCGA 59
DB 98 CTTGAGGTCCTGGGCGGCGGTCGCGGCTCTCCGATGCGCAGATGCTTTAATTCGA 157
QY 60 AAAACCTGGCGGTCGAGGCAAAAAGAGATCTTGGTAAATGGTGTCCAAATCCATCG 119
DB 158 AAAACCTGGCGGTCGAGGCAAAAAGAGATCTTGGTAAATGGTGTCCAAATCCATCG 217
QY 120 CCAACCTTAATAGACAGACAGTAGAGGTGTGTGATGAGCTTACAGAGTACCA 179
DB 218 CCAACCTTAATAGACAGACAGTAGAGGTGTGTGATGAGCTTACAGAGTACCA 277
QY 180 GGGAGTACCCCAAAAAGAGAGGAGAGATCATCAAGACCTTCATCAAGACAG 239
DB 278 GGGAGTACCCCAAAAAGAGAGGAGAGATCATCAAGACCTTCATCAAGACAG 337
QY 240 TCATCAGCTGGGCGGCTCTTAATAGATATAGTATCAAGATGAGTACATGCA 239
DB 338 TCATCAGCTGGGCGGCTCTTAATAGATATAGTATCAAGATGAGTACATGCA 397
QY 300 TGGAGAAATTTAAGAGAAAGTTCATCAGCTGCTATGACCGTGTGATTCATCAG 359
DB 398 TGGAGAAATTTAAGAGAAAGTTCATCAGCTGCTATGACCGTGTGATTCATCAG 457
QY 360 TGGATTAATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAATGAGAGATGC 419
DB 458 TGGATTAATACCTTTGACCGGAATGTGTATCCAGGCTGTAAATGAATGAGAGATGC 517
QY 420 TGGACAAATATTCAGCGGCGGCTCTCTGCAAGTACATGACGGGTAAATATATCT 479
DB 518 TGGACAAATATTCAGCGGCGGCTCTCTGCAAGTACATGACGGGTAAATATATCT 577
QY 480 TTGATCATTTTTCAGATTTGATATTTTGGCTGCTGTAAATCCTTTTGGGAATTTTA 539
DB 578 TTGATCATTTTTCAGATTTGATATTTTGGCTGCTGTAAATCCTTTTGGGAATTTTA 637
QY 540 AACCCCACTTACAAAATCTAT-GTGATGTATCAACAAAATGTTGG 584
DB 638 AACCCCACTTACAAAATCTATGATGTATCAACAAAATGTTGG 683

RESULT 19
AL542926 779 bp mRNA linear EST 24-MAR-2004
LOCUS AL542926 Homo sapiens PLACENTA Homo sapiens cDNA clone CSDB013Y017
DEFINITION 5-PRIME mRNA sequence.
ACCESSION AL542926
VERSION AL542926.3 GI:45718495
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 779)
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:30548565.
Contact: Genoscope

Genoscope - Centre National de Sequençage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 3485.f.
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CSDB013CB090P1c=3485.f.
location/Qualifiers
1.779
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSDB013Y017"
/issue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and EcoRV sites of the
pCMVSPORT 6 vector.
Library was not normalized."

FEATURES

source

ORIGIN

Query Match 92.8%; Score 558; DB 1; Length 779;
Best Local Similarity 97.5%; Pred. No. 1,1e-143;
Matches 552; Conservative 12; Mismatches 2; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCAAAACCTGGCGCTTACGACCAAAAAGAGATCTTGG 95
DB 1 TGGCCACAGATGCTTTAATTCAAAACCTGGCGCTTACGACCAAAAAGAGATCTTGG 60
QY 96 GTAATATGGTGCATATCATGCGACCACTTATAGACACAAAGTAGAGTGC 155
DB 61 GTAATATGGTGCATATCATGCGACCACTTATAGACACAAAGTAGAGTGC 120
QY 156 TGGATGAGCTTACAGAGTACAGGAGATACACCAAAAAGAGAGGCGAGAGA 215
DB 121 TGGATGAGCTTACAGAGTACAGGAGATACACCAAAAAGAGAGGCGAGAGA 180
QY 216 TCATCAAGAACTCATCAAGACATCATCAAGTGCCTTTTATAGAAATATCAT 275
DB 181 TCATCAAGAACTCATCAAGACATCATCAAGTGCCTTTTATAGAAATATCAT 240
QY 276 TTAATCAAGATGAGTACATGATGATGAGAAATTAAGAAAGTTCATGCTGCT 335
DB 241 TTAATCAAGATGAGTACATGATGATGAGAAATTAAGAAAGTTCATGCTGCT 300
QY 336 TGAACGTGTCAGTTTCATCAGTGTGATTAATCACTTGAACCGGAATGTTCAGAGC 395
DB 301 TGAACGTGTCAGTTTCATCAGTGTGATTAATCACTTGAACCGGAATGTTCAGAGC 360
QY 396 TGTAAATGAATGACAGAGATGCTGACCAATTCATGAGCGCACTTCATCTGCAAGT 455
DB 361 TGTAAATGAATGACAGAGATGCTGACCAATTCATGAGCGCACTTCATCTGCAAGT 420
QY 456 CACATGAGAGGGTAAATATGCTTGAATTTTCAATTTTCAATTTTGGCTGCT 515
DB 421 CACATGAGAGGGTAAATATGCTTGAATTTTCAATTTTCAATTTTGGCTGCT 480
QY 516 TGTAAATCCTTTTGGGAATTTTAAACCCCACTTCAAAAAGATGTGATGATCA 575
DB 481 TGTAAATCCTTTTGGGAATTTTAAACCCCACTTCAAAAAGATGTGATGATCA 540
QY 576 AAATGTTGATGAAGAAATATGCA 601
DB 541 AAATGTTGATGAAGAAATATGCA 566

RESULT 20
CR548728

LOCUS CR548728 779 bp mRNA linear EST 12-JUL-2004
 DEFINITION DKFZp469E1232_r1_469 (synonym: pkid1) Pongo pygmaeus cDNA clone
 ACCESSION DKFZp469E1232_5', mRNA sequence.
 VERSION CR548728
 KEYWORDS CR548728.1 GI:50242352
 SOURCE EST.
 ORGANISM Pongo pygmaeus (orangutan)
 Pongo pygmaeus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
 REFERENCE 1 (bases 1 to 779)
 Bahr, A., Lauber, J., Mewes, H.W., Weil, B., Amlid, C., Osanger, A.,
 Fobro, G., Han, M., and Wiemann, S.
 Pongo pygmaeus mRNA (Bahr, A., Lauber, J., Mewes, H.W., et al.)
 JOURNAL Unpublished (2004)
 COMMENT Contact: MIPS
 MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert from S. Wiemann,
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);
 Email: s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen
 (Hilden/Germany) within the cDNA sequencing consortium of the
 German Genome Project. This clone (DKFZp469E1232) is available at
 the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum,
 Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:
 clone@rzpd.de Further information about the clone and the
 sequencing project is available at
 http://mips.gsf.de/projects/cdna/.
 Location/Qualifiers
 1..779
 /organism="Pongo pygmaeus"
 /mol_type="mRNA"
 /db_xref="taxon:9600"
 /clone="DKFZp469E1232"
 /tissue_type="kidney"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_1fb="469 (synonym: pkid1)"
 /note="Vector: pSport1_Sfi; Site_1: SfiIa, Site_2: SfiIb"
 ORIGIN
 Query Match 92.6%; Score 556.4; DB 7; Length 779;
 Best Local Similarity 98.9%; Pred. No. 3e-143;
 Matches 560; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCAAGGACAAAGAGATCTTG 95
 DB 185 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTCAAGGACAAAGAGATCTTG 244
 QY 96 GTAAATGGTGTCATATCCATGCGCCACCTTAATAGACACAAAGTAGAGAGTGC 155
 DB 245 GTAAATGGTGTCATATCCATGCGCCACCTTAATAGACACAAAGTAGAGAGTGC 304
 QY 156 TGGATGAGCTTACAGAGTACCAAGGAGTACACCAAAACAAAGAGAGGAGAGA 215
 DB 305 TGGATGAGCTTACAGAGTACCAAGGAGTACCTCAAAACAAAGAGAGGAGAGA 364
 QY 216 TCATCAAGAACTCATCAAGACAGTCAAGCTGGCCATTTTATAGAAATTAACGT 275
 DB 365 TCATCAAGAACTCATCAAGACAGTCAAGCTGGCCATTTTATAGAAATTAACGT 424
 QY 276 TTAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 335
 DB 425 TTAATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 484
 QY 336 TGAACGTGTCAGTTTCATCAGTGAATTAATCCTTGAACCGAATGTGTTACAGG 395
 DB 485 TGAACGTGTCAGTTTCATCAGTGAATTAATCCTTGAACCGAATGTGTTACAGG 544
 QY 396 TGTAAATGAATCAAGAGATGCTGACCAAAATCATTCAGGCCCACTTCAGCAAGT 455
 DB 545 TGTAAATGAATCAAGAGATGCTGACCAAAATCATTCAGGCCCACTTCAGCAAGT 604

QY 456 CACATGACGGGTTAATATGCTTGCATCAATTTTTCAGATTTGAATTTTGGCTGCT 515
 DB 605 CACATGACGGGTTAATATGCTTGCATCAATTTTTCAGATTTGATTTTGGCTGCT 664
 QY 516 TGTAAATGCTTTTGGGAAATTTTAAACCCCACTTAACAAAACCTATGTGATTCACA 575
 DB 665 TGTAAATGCTTTTGGGAAATTTTAAACCCCACTTAACAAAACCTATGTGATTCACA 724
 QY 576 AATGTGGATGAAGAGAAATATGA 601
 DB 725 AATGTGGATGAAGAGAAATATGA 750
 RESULT 21
 AL558873
 LOCUS 559 bp mRNA linear EST 02-APR-2004
 DEFINITION AL558873 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
 ACCESSION Homo sapiens cDNA clone CS0DJ015YH04 5-PRIME, mRNA sequence.
 VERSION AL558873
 KEYWORDS AL558873.3 GI:46184260
 SOURCE EST.
 ORGANISM Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 559)
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31283006.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 3485.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0DJ015D02Qp1c=3485.f.
 Location/Qualifiers
 1..559
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DJ015YH04"
 /cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
 /cell_line="JURKAT"
 /clone_1fb="Homo sapiens T CELLS (JURKAT CELL LINE) COT
 10-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 ORIGIN
 Query Match 92.3%; Score 554.6; DB 1; Length 559;
 Best Local Similarity 98.6%; Pred. No. 8.8e-143;
 Matches 551; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTACAGGACAAAGAGATCTTG 95
 DB 1 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTACAGGACAAAGAGATCTTG 60
 QY 96 GTAAATGGTGTCATATCCATGCGCCACCTTAATAGACACAAAGTAGAGAGTGC 155
 DB 61 GTAAATGGTGTCATATCCATGCGCCACCTTAATAGACACAAAGTAGAGAGTGC 120
 QY 156 TGGATGAGCTTACAGAGTACCAAGGAGTACCAACCAAAACAAAGAGAGGAGAGA 215
 DB 121 TGGATGAGCTTACAGAGTACCAAGGAGTACCAACCAAAACAAAGAGAGGAGAGA 180

QY 216 TCATCAAGACCTTCATCAAGACAGTGCATCAAGTGGCCATTCTTTATAGGAATTAATAGT 275
 DB 181 TCATCAAGACCTTCATCAAGACAGTGCATCAAGTGGCCATTCTTTATAGGAATTAATAGT 240
 QY 276 TTAATCAAGATGAGCTAGCTATGATGAGAAATTTAAGAAAGTTTCATCAGCTGCTA 335
 DB 241 TTAATCAAGATGAGCTAGCTATGATGAGAAATTTAAGAAAGTTTCATCAGCTGCTA 300
 QY 336 TGACCGTGTGATGATTCATCAGGTGATTAATCCTTTGACCGGAATGTGTATTCAGGC 395
 DB 301 TGACCGTGTGATGATTCATCAGGTGATTAATCCTTTGACCGGAATGTGTATTCAGGC 360
 QY 396 TGTAAATGAAATGAGAGAGATGCTGACCAAAATCATTCAGCCGACCTCATCTGCCAGT 455
 DB 361 TGTAAATGAAATGAGAGAGATGCTGACCAAAATCATTCAGCCGACCTCATCTGCCAGT 420
 QY 456 CACATGAGCGGGTTAAATATGTCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCT 515
 DB 421 CACATGAGCGGGTTAAATATGTCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCT 480
 QY 516 TGTATAATCCTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGATGATCAACA 575
 DB 481 TGTATAATCCTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGATGATCAACA 540
 QY 576 AAATGTTGATGAGAGAA 594
 DB 541 AAATGTTGATGAGAGAA 559

RESULT 22
 LOCUS B1759027 893 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603042814F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:518355 5',
 mRNA sequence.
 ACCESSION B1759027
 VERSION B1759027.1 GI:15750605
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 1 (bases 1 to 893)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: csapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LHAM1457 row: k column: 04
 High quality sequence start: 5
 High quality sequence stop: 869.
 Location/Qualifiers

FEATURES

source
 1..893
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:518355"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_116"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 26 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb.

ORIGIN

Query Match 91.9%; Score 552.4; DB 4; Length 893;
 Best Local Similarity 99.6%; Pred. No. 4e-147; Indels 1; Gaps 1;
 Matches 564; Conservative 0; Mismatches 1;

insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC Library."

QY 36 TGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTTGAGCACAAGAAAGATCTTGG 95
 DB 95 TGGCCACAGATGCTTTAAATCCAAAACCTGGCCGTTGAGCACAAGAAAGATCTTGG 154
 QY 96 GTAAATGCTGTCAAATCCATGCGCACCACTTAAATAGCAGACAAAGTAGTAGGTGC 155
 DB 155 GTAAATGCTGTCAAATCCATGCGCACCACTTAAATAGCAGACAAAGTAGTAGGTGC 214
 QY 156 TGGATGAGCTCTACAGAGTACCAAGGAGTACCCAAAACAAGAGAGGAGAGAGA 215
 DB 215 TGGATGAGCTCTACAGAGTACCAAGGAGTACCCAAAACAAGAGAGGAGAGAGA 274
 QY 216 TCATCAAGAACTCATCAAGACAGTCAATCAAGCTGGCCATTCTTATAGGAATTAATCAGT 275
 DB 275 TCATCAAGAACTCATCAAGACAGTCAATCAAGCTGGCCATTCTTATAGGAATTAATCAGT 334
 QY 276 TTAATCAAGATGAGCTAGCTAGCTAGTGAAGAAATTTAAGAAAGTTTCATCAGCTGCTA 335
 DB 335 TTAATCAAGATGAGCTAGCTAGCTAGTGAAGAAATTTAAGAAAGTTTCATCAGCTGCTA 334
 QY 336 TGACCGTGTGATGATTCATCAGGTGATTAATCCTTTGACCGGAATGTGTATTCAGGC 395
 DB 395 TGACCGTGTGATGATTCATCAGGTGATTAATCCTTTGACCGGAATGTGTATTCAGGC 454
 QY 396 TGTAAATGAAATGAGAGAGATGCTGACCAAAATCATTCAGGCGCACCTCATCTGCCAGT 455
 DB 455 TGTAAATGAAATGAGAGAGATGCTGACCAAAATCATTCAGGCGCACCTCATCTGCCAGT 514
 QY 456 CACATGAGCGGGTTAAATATGTCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCT 515
 DB 515 CACATGAGCGGGTTAAATATGTCTTTGATCATTTTTCAGATTGTAATTTTGGCTGCT 573
 QY 516 TGTATAATCCTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGATGATCAACA 575
 DB 574 TGTATAATCCTTTGGGAATTTTAAACCCCACTTACAAAACATATGTGATGATCAACA 633
 QY 576 AAATGTTGATGAGAGAAATATGTA 601
 DB 634 AAATGTTGATGAGAGAAATATGTA 659

RESULT 23
 LOCUS CN648259 998 bp mRNA linear EST 13-MAY-2004
 DEFINITION ILIDWIGEN_MCO_30108 Katze_MPB Macaca mulatta cDNA clone IBIDW:6795
 5' similar to Bases 72 to 998 highly similar to human TNFAIP8
 (Hs.17839), mRNA sequence.

ACCESSION CN648259
 VERSION CN648259.1 GI:47161702
 KEYWORDS EST.
 SOURCE Macaca mulatta (rhesus monkey)
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Macaca.

REFERENCE
 AUTHORS Katze, M.G., Thomas, M., Korth, M., Iadonato, S.P. and Magness, C.L.
 TITLE Large-scale Rhesus Macaque cDNA Sequencing
 JOURNAL Unpublished (2003)
 COMMENT Contact: C. Magness
 Illumigen Biosciences Inc.
 2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
 Tel: 2063780400

Fax: 2063780408
 Email: cmagnes@illumigen.com
 Sequenced on 2004.03.27. 798 Q20 bases.
 PCR Primers
 FORWARD: CCTCTACTAAGGGAACAA
 BACKWARD: CACTATAGCGCAATGGCTA
 Insert length: 998 Std Error: 0.00
 Plate: CL000256 row: D column: 03
 Seq primer: CCTCTACTAAGGGAACAA
 POLYA=No.

FEATURES

source

Location/Qualifiers
 1. 998
 /organism="Macaca mulatta"
 /mol_type="mRNA"
 /strain="Indian"
 /db_xref="taxon:9544"
 /clone="IBIDW:6795"
 /sex="male"
 /tissue_type="blood"
 /cell_type="PBMC"
 /dev_stage="adult"
 /lab_host="E. coli: SOLR"
 /clone_lib="Katzel MMPB"
 /note="Vector: Uni-ZAP XR, Site 1: EcoR I, Site 2: Xho I;
 Created from Stratagene ZAP-CDNA Synthesis kit (Catalog
 #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit
 (Catalog #200450)"

ORIGIN

Query Match 91.2%; Score 548.4; DB 7; Length 998;
 Best Local Similarity 98.1%; Pred. No. 5.4e-141;

Matches 555; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

36 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTACGGACACAAAGAAATCTTGG 95
 |||||
 94 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTACGGACACAAAGAAATCTTGG 153
 |||||
 96 GTPAAATGCTGCCAATCATCGCCACCACTTAATAGACGACAGAGTGAAGTGC 155
 |||||
 154 GTPAAATGCTGCCAATCATCGCCACCACTTAATAGACGAGTGAAGTGC 213
 |||||
 156 TGGATGAGCTTACAGAGTGACCGAGAGTACACCCAAACAGAGAGGACAGAGA 215
 |||||
 214 TGGATGAGCTTACAGAGTGACCGAGAGTACACCCAAACAGAGAGGACAGAGA 273
 |||||
 216 TCATCAAGAACTCATCAAGAGTCAAGTGGCCATTTTATAGGAATATCACT 275
 |||||
 274 TCATCAAGAACTCATCAAGAGTCAAGTGGCCATTTTATAGGAATATCACT 333
 |||||
 276 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTCACTGCTGCTA 335
 |||||
 334 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTCACTGCTGCTA 393
 |||||
 336 TGAACGTGCTCACTTCCATCAAGTGAATTAATCTTGAACCGAATGTGTTATCCAGGC 395
 |||||
 394 TGAACGTGCTCACTTCCATCAAGTGAATTAATCTTGAACCGAATGTGTTATCCAGGC 453
 |||||
 396 TGTAAATGAATGCAAGAGATGCTGCAACAAATTCATTCAGCGCCACTCACTGCCAAGT 455
 |||||
 454 TGTAAATGAATGCAAGAGATGCTGCAACAAATTCATTCAGCGCTCACTCACTGCCAAGT 513
 |||||
 456 CACATGAGCGGTTAATTAATGCTTTGATCAATTTTCAAGTTGTGAATTTTGGCTGCT 515
 |||||
 514 CACATGAGCGGTTAATTAATGCTTTGATCAATTTTCAAGTTGTGAATTTTGGCTGCT 573
 |||||
 516 TGTAAATGCTTTTGGAAATTTTAAACCCACTTACAAAACTATGTGATGATCAACA 575
 |||||
 574 TGTAAATGCTTTTGGAAATTTTAAACCTTACAAAACTATGTGATGATGATCAACA 633
 |||||
 576 AAATGTTGATGAAGAACTATGA 601
 |||||
 634 AAATGTTGATGAAGAACTATGA 659
 |||||

RESULT 24

AUI35377

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AUI35377 816 bp mRNA linear EST 02-ANG-2002
 AUI35377 PLACE1 Homo sapiens CDNA clone PLACE1001920 5', mRNA
 sequence.
 AUI35377
 AUI35377.1 GI:10995916
 EST.
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 Ota.T., Nishikawa.T., Suzuki.Y., Ishii.S., Saio.K., Kawai.Y.,
 Yamamoto.J., Wakamatsu.A., Nakamura.Y., Nagai.T., Sugano.S. and
 Isogai.T.
 HRI human CDNA project
 Unpublished (2000)
 Contact: Takao Isogai
 Genomics Laboratory
 Helix Research Institute
 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
 Tel: 81-438-52-3975
 Fax: 81-438-52-3986
 Email: genomics@ri.co.jp
 HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix
 Research Institute; CDNA library construction; Department of
 Virology, Institute of Medical Science, University of Tokyo, and
 Helix Research Institute.
 Location/Qualifiers
 1. 816
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="PLACE1001920"
 /tissue_type="Placenta"
 /clone_lib="PLACE1"
 /note="Vector: pME18SFL3"

FEATURES

source

Query Match 91.1%; Score 547.6; DB 1; Length 816;
 Best Local Similarity 98.9%; Pred. No. 8.5e-141;
 Matches 562; Conservative 0; Mismatches 4; Indels 2; Gaps 1;

ORIGIN

36 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTACGGACACAAAGAAATCTTGG 95
 |||||
 102 TGGCCACAGATGCTTTAATTCGAAACCTGGCCGTTACGGACACAAAGAAATCTTGG 161
 |||||
 96 GTPAAATGCTGCCAATCATCGCCACCACTTAATAGACGACAGAGTGAAGTGC 155
 |||||
 162 GTPAAATGCTGCCAATCATCGCCACCACTTAATAGACGACAGAGTGAAGTGC 221
 |||||
 156 TGGATGAGCTTACAGAGTGACCGAGAGTACACCCAAACAGAGAGGACAGAGA 215
 |||||
 222 TGGATGAGCTTACAGAGTGACCGAGAGTACACCCAAACAGAGAGGACAGAGA 281
 |||||
 216 TCATCAAGAACTCATCAAGAGTCAAGTGGCCATTTTATAGGAATATCACT 275
 |||||
 282 TCATCAAGAACTCATCAAGAGTCAAGTGGCCATTTTATAGGAATATCACT 341
 |||||
 276 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTCACTGCTGCTA 335
 |||||
 342 TTAATCAAGATGAGCTAGCATTTGATGAGAAATTTAAGAAAGATTCACTGCTGCTA 401
 |||||
 336 TGAACGTGCTCACTTCCATCAAGTGAATTAATCTTGAACCGAATGTGTTATCCAGGC 395
 |||||
 402 TGAACGTGCTCACTTCCATCAAGTGAATTAATCTTGAACCGAATGTGTTATCCAGGC 461
 |||||
 396 TGTAAATGAATGCAAGAGATGCTGCAACAAATTCATTCAGCGCCACTCACTGCCAAGT 455
 |||||
 462 TGTAAATGAATGCAAGAGATGCTGCAACAAATTCATTCAGCGCCACTCACTGCCAAGT 521
 |||||

Query Match 90.2%; Score 542.4; DB 6; Length 596;
 Best Local Similarity 99.6%; Pred. No. 2,2e-139;
 Matches 554; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 CTTACGCGTCCGCGCGCGCGTCCG-CGACTCTCCGATGGCCAGATGCTTTAATTCGA 59
    |||
DB 40 CTTACGCGTCCGCGCGCGCGCGCGCTCCGATGGCCAGATGCTTTAATTCGA 99
    |||
QY 60 AAAACCTGCGCGCTTCAGGCAAAAAGATCTTGGGTAAATGCTGCCAATTCATCG 119
    |||
DB 100 AAAACCTGCGCGCTTCAGGCAAAAAGATCTTGGGTAAATGCTGCCAATTCATCG 159
    |||
QY 120 CCACCACTTTAATAGACGACACAAGTAGTGGTGTGCTGAGTACTCTACAGAGTACCA 179
    |||
DB 160 CCACCACTTTAATAGACGACACAAGTAGTGGTGTGCTGAGTACTCTACAGAGTACCA 219
    |||
QY 180 GGGAGTAGACCCCAAAACAAGAGAGGACAGAAAGATCATCAAGAACTTCATCAAGACAG 239
    |||
DB 220 GGGAGTAGACCCCAAAACAAGAGAGGACAGAAAGATCATCAAGAACTTCATCAAGACAG 279
    |||
QY 240 TCATCAAGCTGGCGCTTTTATAGGAATATCATGTTTATCAAGTAGCTAGCTTGA 299
    |||
DB 280 TCATCAAGCTGGCGCTTTTATAGGAATATCATGTTTATCAAGTAGCTAGCTTGA 339
    |||
QY 300 TGGAGAAATTTAAGAGAAAGTTCATCAGCTGCTANTGACCGTGTCAAGTTTCATCAG 359
    |||
DB 340 TGGAGAAATTTAAGAGAAAGTTCATCAGCTGCTANTGACCGTGTCAAGTTTCATCAG 399
    |||
QY 360 TGGATTATACCTTTGACCGGAATGTGTATTCAGGCTGTTAAATGAGTAGAGAGATGC 419
    |||
DB 400 TGGATTATACCTTTGACCGGAATGTGTATTCAGGCTGTTAAATGAGTAGAGAGATGC 459
    |||
QY 420 TGCACCAATCATTCAGCGCCACCTCATGCGCAAGTCAATGACGGGTTAAATGAGTCT 479
    |||
DB 460 TGCACCAATCATTCAGCGCCACCTCATGCGCAAGTCAATGACGGGTTAAATGAGTCT 519
    |||
QY 480 TGGATCATTTTTCAGATTTGATTTTGGCTGCTGTTATATCTTTTGGGAATTTTA 539
    |||
DB 520 TGGATCATTTTTCAGATTTGATTTTGGCTGCTGTTATATCTTTTGGGAATTTTA 579
    |||
QY 540 AACCCCACTTACAAA 555
    |||
DB 580 AACCCCACTTACAAA 595
    |||

```

RESULT 27
 BX405197 834 bp mRNA linear EST 03-MAY-2004
 LOCUS BX405197 Homo sapiens T CELLS (JURKAT CELL LINE) Homo sapiens cDNA
 DEFINITION clone CS0DH001YB17 5-PRIME, mRNA sequence.
 ACCESSION BX405197
 VERSION BX405197.2 GI:46952678
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 834)
 Li,W.B., Gruber,C., Jessee,J. and Polyes,D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 15, 2003 this sequence version replaced gi:30762567.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 3485.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna7e=CS0AH001CA09Qp1kc=3485.f>.
 FEATURES
 Location/Qualifiers
 1..834
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DH001YB17"
 /tissue_type="T CELLS (JURKAT CELL LINE)"
 /cell_line="JURKAT CELL LINE"
 /clone_1ib="Homo sapiens T CELLS (JURKAT CELL LINE)"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Query Match 89.6%; Score 538.6; DB 5; Length 834;
 Best Local Similarity 97.4%; Pred. No. 2.7e-138;
 Matches 552; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

```

QY 36 TGGCCACAGATGCTTTAATTCAAAACCTGCGCGCTTCAGGACACAAAAGATCTTGG 95
    |||
DB 1 TGGCCACAGATGCTTTAATTCAAAACCTGCGCGCTTCAGGACACAAAAGATCTTGG 60
    |||
QY 96 GTAAATGGTGTCCAAATTCATGCGCACCACTTAATGAGACACAAGTAGTGGTGC 155
    |||
DB 61 GTAAATGGTGTCCAAATTCATGCGSACCCTTAATGAGACACAAGTAGTGGTGC 120
    |||
QY 156 TGGATGAGCTCTTACAGAGTACAGCAGGAGTACACCAAAACAAGAGAGGACAGAGA 215
    |||
DB 121 TGGATGAGCTCTTACAGAGTACAGCAGGAGTACACCAAAACAAGAGAGGACAGAGA 180
    |||
QY 216 TCATCAAGAACTCTCAACAGACATCATCAGCTGCCATTTTATAGGAATATCACT 275
    |||
DB 181 TCATCAAGAACTCTCAACAGACATCATCAGCTGCCATTTTATAGGAATATCACT 240
    |||
QY 276 TTAATCAAGATAGCTAGATGATGAGAAATTTAAGAGAAAGTCAATCAGCTTGCTA 335
    |||
DB 241 TTAATCAAGATAGCTAGATGATGAGAAATTTAAGAGAAAGTCAATCAGCTTGCTA 300
    |||
QY 336 TGACGCTGTGCTAGTTTCCATCAGTGTGATTATACC-TTTCACCGGAATGTATTACAG 394
    |||
DB 301 TGACGCTGTGCTAGTTTCCATCAGTGTGATTATACC-TTTCACCGGAATGTATTACAG 360
    |||
QY 395 CTGTTAATGAATGACAGAGATGCTGCACCAATCATCAGCGCCACTCATCTGCCAAG 454
    |||
DB 361 CTGTTAATGAATGACAGAGATGCTGCACCAATCATCAGCGCGAGCTCATCTGCCAAG 420
    |||
QY 455 TCACATGACGCGGTTAATTAATGCTTTGATCAATTTTTCAGATTTGGAATTTTGGCTGCC 514
    |||
DB 421 TCACATGACGCGGTTAATTAATGCTTTGATCAATTTTTCAGATTTGGAATTTTGGCTGCC 480
    |||
QY 515 TTGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGATATCAAC 574
    |||
DB 481 TTGTATATCCTTTTGGGAATTTTAAACCCCACTTACAAAACCTATGATGATATCAAC 540
    |||
QY 575 AAAATGTTGATGAAGAAGAACTATGTA 601
    |||
DB 541 AAAATGTTGATGAAGAAGAACTATGTA 567
    |||

```

RESULT 28
 CDS20304 747 bp mRNA linear EST 06-JUN-2003
 LOCUS CDS20304
 DEFINITION AGENCOURT 14360208 NIH MGC 191 Homo sapiens cDNA clone
 IMAGE:30410510 5', mRNA sequence.
 ACCESSION CDS20304
 VERSION CDS20304.1 GI:31452022
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: csabds-remail.nih.gov
 Tissue Procurement: Narayan Bhat
 CDNA Library Preparation: Clontech Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 Plate: NDCM200 row: m column: 15
 High quality sequence stop: 499.
 Location/Qualifiers

FEATURES

source

1. 747
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30410510"
 /tissue_type="Pooled"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_id="NIH MGC 191"
 /note="Vector: pDNR-LIB; Site 1: SfiI (ggccatratggcc); Site 2: SfiI (ggccgctggcc); Library is oligo-dT primed and directionally cloned. PBMC - Peripheral Blood Mononuclear Cells. RNA was pooled from 3/6hour stimulation with PMA adn Ionomycin. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGSCC-3' and 3' adaptor sequence: 5'-ATTCTAAGAGCCGAGGCGCCGACATG-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.69 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 89.6%; Score 538.4; DB 6; Length 747;
 Best Local Similarity 97.2%; Pred. No. 3e-138;
 Matches 548; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
 Db 36 TGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTGAGCAGCAAAAGAGATCTTGG 95
 107 TGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTGAGCAGCAAAAGAGATCTTGG 166
 Qy 96 GTAAATGCTGTCCTAATTCAGATGCGCCACCACTTAATAGACACAAAGTAGAGGTGC 155
 Db 167 GTAAATGCTGTCCTAATTCAGATGCGCCACCACTTAATAGACACAAAGTAGAGGTGC 226
 Qy 156 TGGATGAGCTCTACAGAGTACAGAGAGTACACCAAAACAAGAGAGAGAGAGA 215
 Db 227 TGGATGAGCTCTACAGAGTACAGAGAGTACACCAAAACAAGAGAGAGAGAGA 286
 Qy 216 TCATCAAGAACTCATTAAGACAGTCTCAAGTGGCCATTTTATAGAAATATCAGT 275
 Db 287 TCATCAAGAACTCATTAAGACAGTCTCAAGTGGCCATTTTATAGAAATATCAGT 346
 Qy 276 TTATCAAGATGAGCTGAGATTTGAAGAAATTTAAGAGAAAGTTCATCAGCTGCTA 335
 Db 347 TTATCAAGATGAGCTGAGATTTGAAGAAATTTAAGAGAAAGTTCATCAGCTGCTA 406
 Qy 336 TGAACGCTGCTGATTCATCAGGTGATTAATCCTTTGACCGAATGTGTATCCAGGC 395
 Db 407 TGAACGCTGCTGATTCATCAGGTGATTAATCCTTTGACCGAATGTGTATCCAGGC 466
 Qy 396 TGTAAATGATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCAGTCCAGT 455

Db 467 TGTAAATGATGACAGAGATGCTGACCAAAATCATTCAGCGCCACCTCAGTCCAGT 526
 Qy 456 CACATGAGAGGCTTAATTAATGCTTTCATTCATTTTCAGATTTGATTTGGTGGCT 515
 Db 527 CACATGAGAGGCTTAATTAATGCTTTCATTCATTTTCAGATTTGATTTGGTGGCT 566
 Qy 516 TGTAAATGCTTTTGGGAAATTTTAAACCCCACTTACAAAACATGATGATGATCAACA 575
 Db 587 TGTAAATGCTTTTGGGAAATTTTAAACCCCACTTACAAAACATGATGATGATGATCAACA 646
 Qy 576 AATGTTGATGAGAGACATAT 599
 Db 647 AATGTTGATGAGAGACATAT 670

RESULT 29

LOCUS CD701806 622 bp mRNA linear EST 25-JUN-2003
 DEFINITION EST18330 human nasopharynx Homo sapiens cDNA, mRNA sequence.
 ACCESSION CD701806
 VERSION CD701806.1 GI:32232436
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 622)

Liou X.-X., Zhou, Y., Zhang, L.-J., Xu, H., Chen, H.-K., Pan, Z.-G. and
 Transcriptional Gene Expression Profile of Human Nasopharynx
 Unpublished (2003)
 Contact: Yixin Zeng
 Cancer Center
 Sun Yat-sen University
 651 Dongfeng Road East, Guangzhou 510060, China
 Tel: 86-1380-9770-743
 Fax: 86-20-8775-4506
 Email: yxzeng@zsums.edu.cn.

FEATURES

source

1. 622
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="normal nasopharynx"
 /clone_id="human nasopharynx"
 /note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN

Query Match 89.4%; Score 537; DB 6; Length 622;
 Best Local Similarity 100.0%; Pred. No. 7e-138;
 Matches 537; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 36 TGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTGAGCAGCAAAAGAGATCTTGG 95
 Qy 86 TGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTGAGCAGCAAAAGAGATCTTGG 145
 Db 96 GTAAATGCTGTCCTAATTCAGATGCGCCACCACTTAATAGACACAAAGTAGAGGTGC 155
 Qy 146 GTAAATGCTGTCCTAATTCAGATGCGCCACCACTTAATAGACACAAAGTAGAGGTGC 205
 Db 156 TGGATGAGCTCTACAGAGTACAGAGAGTACACCAAAACAAGAGAGAGAGAGA 215
 Qy 206 TGGATGAGCTCTACAGAGTACAGAGAGTACACCAAAACAAGAGAGAGAGAGA 265
 Db 216 TCATCAAGAACTCATTAAGACAGTCTCAAGTGGCCATTTTATAGAAATATCAGT 275
 Qy 266 TCATCAAGAACTCATTAAGACAGTCTCAAGTGGCCATTTTATAGAAATATCAGT 325
 Db 276 TTAAATCAAGATGAGCTGAGATTTGAAGAAATTTAAGAGAAAGTTCATCAGCTGCTA 335
 Qy 326 TTAAATCAAGATGAGCTGAGATTTGAAGAAATTTAAGAGAAAGTTCATCAGCTGCTA 385

QY 336 TGACCGTGTGAGTTTCCATGAGTGATTTATACCTTTGACCGGGAATGTTATCCAGGC 395
 DB 386 TGACCGTGTGAGTTTCCATGAGTGATTTATACCTTTGACCGGGAATGTTATCCAGGC 445
 QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGCCACCTCATTCCCAAGT 455
 DB 446 TGTAAATGAATGACAGAGATGCTGCACCAATTCATTGACGCCACCTCATTCCCAAGT 505
 QY 456 CACATGAGCGGGTTAATTAATGCTTTGATCATTTTTCAGATGTTGAATTTTGGCTGCT 515
 DB 506 CACATGAGCGGGTTAATTAATGCTTTGATCATTTTTCAGATGTTGAATTTTGGCTGCT 565
 QY 516 TGTATATGCTTTTGGGAATTTTAAACCCACTTACAAAACATGATGATGATATCA 572
 DB 566 TGTATATGCTTTTGGGAATTTTAAACCCACTTACAAAACATGATGATGATATCA 622

RESULT 30
 BP228279 580 bp mRNA linear EST 15-SEP-2004
 LOCUS BP228279 Sugano cDNA library, dermoid cancer Homo sapiens cDNA
 DEFINITION Clome DMC08945, mRNA sequence.
 ACCESSION BP228279
 VERSION BP228279.1 GI:52101189
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 580)
 Suzuki Y., Yamashita R., Shirota M., Sakakibara Y., Chiba J., Mizushima-Sugano J., Nakai K. and Sugano S. Sequence comparison of human and mouse genes reveals a homologous block structure in the promoter regions
 Genome Res. 14 (9), 1711-1718 (2004)
 CONTACT: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.
 Location/Qualifiers
 1..580
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="DMC08945"
 /clone_lib="Sugano cDNA library, dermoid cancer"
 /note="dermoid cancer"

ORIGIN

Query Match 89.0%; Score 535, DB 5, Length 580;
 Best Local Similarity 99.6%; Pred. No. 2.5e-137;
 Matches 546; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 54 ATTGCAAAAACCGGCGCTTCCAGGCAAAAAGATCTGGTAAATGCTCCAAAT 113
 DB 1 ATTGCAAAAACCGGCGCTTCCAGGCAAAAAGATCTGGTAAATGCTCCAAAT 59
 QY 114 CCATGCGCACCACTTATATGACGACACAAGTAGAGTGCTGATGAGCTTACAGAG 173
 DB 60 CCATGCGCACCACTTATATGACGACACAAGTAGAGTGCTGATGAGCTTACAGAG 119
 QY 174 TGACCGGAGATGACCCCAAAACAGAGAGGCGAGAAAGATCATCAGACCTCATCA 233
 DB 120 TGACCGGAGATGACCCCAAAACAGAGAGGCGAGAAAGATCATCAGACCTCATCA 179
 QY 234 AGACATCATCAAGCTGGCATCTTATATGAAATATCATGTTATCAAGATGACTAG 293
 DB 180 AGACATCATCAAGCTGGCATCTTATATGAAATATCATGTTATCAAGATGACTAG 239
 QY 294 CATGTATGGAATTTAAGAAAGATTCATCAGCTTGCTATGACCGTGTGCTTCC 353

DB 240 CATGTATGAGAAATTTAAGAAAGATTCATCAGCTTGCTATGACCGTGTGCTTCC 299
 QY 354 ATCAGGTGATTTATACCTTTGACCGGAATGTTATTCAGGCTGTTAAATGAATCAGAG 413
 DB 300 ATCAGGTGATTTATACCTTTGACCGGAATGTTATTCAGGCTGTTAAATGAATCAGAG 359
 QY 414 AGATGTGACCAATCATTTAGCGGCCACCTCAGTCGCAAGTCATGAGACGGGTTAATA 473
 DB 360 AGATGTGACCAATCATTTAGCGGCCACCTCAGTCGCAAGTCATGAGACGGGTTAATA 419
 QY 474 ATGCTCTTGATCATTTTTCAGATGTTGAATTTTGGCTGCTTATATACTTTTGGGA 533
 DB 420 ATGCTCTTGATCATTTTTCAGATGTTGAATTTTGGCTGCTTATATACTTTTGGGA 479
 QY 534 ATTTTAAACCCCACTTACAAAACATGATGATGATGATCAAAAATGTTGATGAAGAGA 593
 DB 480 ATTTTAAACCCCACTTACAAAACATGATGATGATGATCAAAAATGTTGATGAAGAGA 539
 QY 594 ACATATGA 601
 DB 540 ACATATGA 547

RESULT 31
 CN305328 672 bp mRNA linear EST 16-MAY-2004
 LOCUS CN305328 17000599947340 GRN_PRENEM Homo sapiens cDNA 5', mRNA sequence.
 DEFINITION CN305328
 ACCESSION CN305328
 VERSION CN305328.1 GI:47321742
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 672)
 Brandenberger R., Wei H., Zhang S., Lei S., Murage J., Flisk G.J., Li Y., Xu C., Fang R., Guejler K., Rao M.S., Mandalam R., Lebkowski J. and Stanton L.W. Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
 Nat. Biotechnol. 22 (6), 707-716 (2004)
 CONTACT: Brandenberger R
 Regenerative Medicine
 Geron Corporation
 230 Constitution Drive, Menlo Park, CA 94025, USA
 Tel: 650 473 8658
 Fax: 650 473 7760
 Email: rbrandenberger@geron.com
 Insert Length: 672 Std Error: 0.00.
 Location/Qualifiers
 1..672
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /ligase_type="embryonic stem cell, retinoic acid and mitogen-treated hES cell line H7"
 /clone_lib="GRN_PRENEM"
 /note="oligo dt primed, full-length enriched cDNA library from hES cell line H7 (p29) maintained in feeder-free conditions. Embryoid bodies were generated in the presence of all-trans retinoic acid and mitogens."

ORIGIN

Query Match 88.8%; Score 533.6, DB 7, Length 672;
 Best Local Similarity 98.8%; Pred. No. 6.3e-137;
 Matches 569; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY 1 CTTGACGCTCCGGGCGCTGCG-CGACTCTTCGATGCGCACATGCTTAAATCCA 59
 DB 97 CTTGACGCTCCGGGCGCTGCGCACCTCCGATGCGCACAGATGCTTAAATCCA 156
 QY 60 AAAACGTGCGCTTACGCAAAAAGATCTTGGGTAAATGTTGTCAAATCATCG 119

Db 157 AAAAAGCTGGCGTTGAGGACCAAAAAGAGATCTGGTAAATATGTTGCCAAATCCATCG 216

Qy 120 CCACCACTTATATGAGACACAACTGTGAGGTGCTGATGAGCTCTTACAGAGTGACCA 179

Db 217 CCACCACTTATATGAGACACAACTGTGAGGTGCTGATGAGCTCTTACAGAGTGACCA 276

Qy 180 GGGAGTACACCCCAAAAACAAGAGGACAGAGAGATCATCAAGAACCTCATCAAGACAG 239

Db 277 GGGAGTACACCCCAAAAACAAGAGGACAGAGAGATCATCAAGAACCTCATCAAGACAG 336

Qy 240 TCATCAAGTGGCCATCTTATAGAAATATATCATGTTAATCAAGATGAGTACATTTGA 299

Db 337 TCATCAAGTGGCCATCTTATAGAAATATATCATGTTAATCAAGATGAGTACATTTGA 396

Qy 300 TGAAGAAATTTAAGAGAAAGTTTCATGACCTGCTATGACCGTGTGATGATTTCCATGAG 359

Db 397 TGAAGAAATTTAAGAGAAAGTTTCATGACCTGCTATGACCGTGTGATGATTTCCATGAG 456

Qy 360 TGAATTTATACCTTTGACCGGAATGTGTATCCAGGCTTTAATATGATGACAGAGATGC 419

Db 457 TGAATTTATACCTTTGACCGGAATGTGTATCCAGGCTTTAATATGATGACAGAGATGC 516

Qy 420 TGCACCAATATATTCAGCGCCACCTCATGCTCCAAAGTACAT- GGAAGGTTAATATATGTC 478

Db 517 TGCACCAATATATTCAGCGCCACCTCATGCTCCAAAGTACATGGAAGGTTAATATATGTC 576

Qy 479 TTTGATCATTTTTCAGATTTGTAATTTTGGCTGCTGTATAT- CTTTGGGAATTT 537

Db 577 GTTATCATTTTTCAGATTTGTAATTTTGGCTGCTGTATATCCTTTTGGGAATTT 636

Qy 538 TAAACCCCACTTACAAAACCTATGTATGTATGATCA 573

Db 637 TAAACCCCACTTACAAAACCTATGTATGTATGATCA 672

RESULT 32

LOCUS CV023086 566 bp mRNA linear EST 20-AUG-2004

DEFINITION 5463 Full length cDNA from the Mammalian Gene Collection Homo sapiens cDNA 5' similar to BC007014, mRNA sequence.

ACCESSION CV023086

VERSION CV023086.1 GI:51480836

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 566)

AUTHORS Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S., Dricot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O., Clingingsmith,T.R., Hartley,J.L., Espocho,D., Cheo,D., Moore,T., Slimons,B., Segueria,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C., Vandenhaute,J., Quisic,M.E., Alcala,J.S., Hill,D.E. and Vidal,M.

TITLE Human ORFome Version 1.1: a Platform for Reverse Proteomics

JOURNAL Genome Res. (2004) In press

COMMENT Contact: Vidal M

Marc Vidal Laboratory

Dana Farber Cancer Institute

1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA

Tel: 617 632 5180

Fax: 617 632 5739

Email: Marc.Vidal@dfci.harvard.edu

ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF results from a PCR reaction using an MGC full-length cDNA as template DNA and ORF specific primers

PCR Primers

FORWARD: ATGCACTCCGAGACAGAGAATCCAA

BACKWARD: CATATGTTCTTTCATCAACATTT

Insert length: 566 Std Error: 21.00

Plate: 11075 row: 05 column: F

Seq primer: ACTGCGCGTGTTCACAGCGTGTGACGTGAGGAAAC

High quality sequence start: 104

High quality sequence stop: 565

FEATURES

POLYA=NO.

Location/Qualifiers

1..566

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="mixed"

/clone_lib="Full Length cDNA from the Mammalian Gene Collection"

/note="Vector: mixed; The ORFs were PCR amplified from the MGC (Mammalian Gene Collection) as of April 2004 and cloned by recombinational Gateway cloning into pDONR223 Donor vector. Reference : MGC (Mammalian Gene Collection) Program Team, Generation and Initial Analysis of more than 15,000 Full-length Human and Mouse cDNA Sequences. PNAS, 2002, 99(126), 16899-16903"

ORIGIN

Query Match 88.8%; Score 533.4; DB 7; Length 566;

Best Local Similarity 99.8%; Pred. No. 6.8e-137;

Matches 534; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 36 TGGCCACAGATGTCCTTAAATTCAAAACCTGGCCGTCAGGACAAAAGAATCTTGG 95

Db 32 TGGCCACAGATGTCCTTAAATTCAAAACCTGGCCGTCAGGACAAAAGAATCTTGG 91

Qy 96 GTAAATGTCGTCCAAATCATGCGCACCACTTAAATAGACACAAAGTATGAGGTGC 155

Db 92 GTAAATGTCGTCCAAATCATGCGCACCACTTAAATAGACACAAAGTATGAGGTGC 151

Qy 156 TGGATGAGCTTACAGAGTACACAGAGATGACCCAAAACAAGAGAGGACAGAGA 215

Db 152 TGGATGAGCTTACAGAGTACACAGAGATGACCCAAAACAAGAGAGGACAGAGA 211

Qy 216 TCATCAAGAACCTCATCAAGACGATGACAGTGGCCATCTTTATAGAAATATCAT 275

Db 212 TCATCAAGAACCTCATCAAGACGATGACAGTGGCCATCTTTATAGAAATATCAT 271

Qy 276 TTAATCAAGTATGAGTATGATGATGAGAAATTTAAGAAAGATTCATCAGTGTCT 335

Db 272 TTAATCAAGTATGAGTATGATGATGAGAAATTTAAGAAAGATTCATCAGTGTCT 331

Qy 336 TGAACGTCGTCACTTCATCAGGTGATTAATCTTTGACCGGAATGTATTCAGGC 355

Db 332 TGAACGTCGTCACTTCATCAGGTGATTAATCTTTGACCGGAATGTATTCAGGC 391

Qy 396 TGTAAATGAATGACAGAGATGCTGACCAATCATTTAGGCCCACTTACGCCAAT 455

Db 392 TGTAAATGAATGACAGAGATGCTGACCAATCATTTAGGCCCACTTACGCCAAT 451

Qy 456 CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATTTGTAATTTTGGCTGCT 515

Db 452 CACATGACGGGTTAATATGCTTTGATCATTTTTCAGATTTGTAATTTTGGCTGCT 511

Qy 516 TGTATATCTTTTGGAAATTTTAAACCCCATTAACAAAACATATGTATGTAT 570

Db 512 TGTATATCTTTTGGAAATTTTAAACCCCATTAACAAAACATATGTATGTAT 566

RESULT 33

LOCUS CR557522 707 bp mRNA linear EST 13-JUL-2004

DEFINITION DKFZ0469E2119 r1 469 (synonym: pkid1) Pongo pygmaeus cDNA clone

ACCESSION CR557522

VERSION CR557522.1 GI:50278801

KEYWORDS EST.

SOURCE Pongo pygmaeus (orangutan)

ORGANISM Pongo pygmaeus

REFERENCE 1 (bases 1 to 707)

AUTHORS Pouetka,A., Albert,R., Moosmayer,P., Schnupp,I., Wellenreuther,R.,

TITLE
Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.
JOURNAL
Pongo pygmaeus mRNA (Poustka,A., Albert,R., Moosmayer,P., et al.)
COMMENT
Unpublished (2004)
Contact: MIPS
MIPS

Ingolstaedter Landstr.1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; any. Please contact RZPD for ordering:
http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=DKFZp49E2119
Further information about the clone and the sequencing project is available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
location/Qualifiers

1..707
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp49E2119"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1ib="469 (synonym: pkid1)"
/note="Vector: pSPORT1_sfi; site_1: sfi1A; site_2: sfi1B"

ORIGIN

Query Match 88.3%; Score 530.6; DB 7; Length 707;
Best Local Similarity 99.3%; Pred. No. 4.3e-136;
Matches 533; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

QY 36 TGGCCACAGATGCTTAATTCACAAAACCTGGCGCTTACAGGACAAAAGAGATCTTG 95
DB 171 TGGCCACAGATGCTTAATTCACAAAACCTGGCGCTTACAGGACAAAAGAGATCTTG 230
QY 96 GTAAATGCTGTCACAAATCCATCGCCACGACCTTAATAGACACACAGTAGTAGTGC 155
DB 231 GTAAATGCTGTCACAAATCCATCGCCACGACCTTAATAGACACAGTAGTAGTGC 290
QY 156 TGGAGAGCTTCACAGAGTACACAGGAGTATCACCCAAAAGAGAGGACAGAGAA 215
DB 291 TGGAGAGCTTCACAGAGTACACAGGAGTATCACCCAAAAGAGAGGACAGAGAA 350
QY 216 TCATCAGAACCTCATCAAGACAGTATCAAGCTGCGCACTTTATAGGAATATCACT 275
DB 351 TCATCAGAACCTCATCAAGACAGTATCAAGCTGCGCACTTTATAGGAATATCACT 410
QY 276 TTAATCAGATGAGTACGATGATGAGAGAAATTTAAGAGAAAGTTCACTGCTCTA 335
DB 411 TTAATCAGATGAGTACGATGATGAGAGAAATTTAAGAGAAAGTTCACTGCTCTA 470
QY 336 TGAACCTGCTCACTTCCATCAGTGTGATTTATACCCGGAATGTGTATCCAGGC 395
DB 471 TGAACCTGCTCACTTCCATCAGTGTGATTTATACCCGGAATGTGTATCCAGGC 530
QY 396 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGGCGCACTCATCGCAAGT 455
DB 531 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGGCGCTCATCGCAAGT 590
QY 456 CACATGACGGGTAAATTAATGCTTTGATCATTTTTCAGATGTGAATTTTGGCTGCT 515
DB 591 CACATGACGGGTAAATTAATGCTTTGATCATTTTTCAGATGTGAATTTTGGCTGCT 650
QY 516 TGTATTAATCTTTTGGGAATTTTAAACCCCATTAACAAAACCTATGTATGATGATCA 572
DB 651 TGTATTAATCTTTTGGGAATTTTAAACCCCATTAACAAAACCTATGTATGATGATCA 707

```

RESULT 34
CK903031 572 bp mRNA 1linear EST 11-MAR-2004
LOCUS CK903031
DEFINITION 1155c03.x5 HR85 islet Homo sapiens cDNA clone IMAGE:6548934 3'
similar to TR:095379 O95379 MDC-3.13 ISOFORM 2. [2] TR:Q9UP47 ;

ACCESSION
CK903031
VERSION
CK903031.1 GI:45364562
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE
AUTHORS
Wyllie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarelisvili,R., Williams,T., Jackson,Y. and Bowers,Y.

TITLE
WashU-Harvard Pancreas EST Project
JOURNAL
Unpublished (2000)
COMMENT
Other ESTs: 1155c03.y1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
This read is a 3' RESEQUENCE of a previously sequenced pancreas clone

Good hit to opposite strand read. . . wrong orientation BUT PASSED FOR
MOUSE-PANCREAS VERIFICATION
Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco.
Location/Qualifiers

FEATURES
source

1..572
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6548934"
/tissue_type="Purified pancreatic islet"
/lab_host="DH10B"
/clone_1ib="HR85 islet"
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1: NotI; Site 2: XhoI; cDNA made by oligo-dT priming. Size-selected on agarose gel. Average insert size -1kb. 5' xhoI site was destroyed after directional cloning. Amplified once. Contact information: Hiroshi Inoue, MD, Metabolism Div. (Alan Reimut Lab), Washington University School of Medicine, Box 8127, 660 South Euclid Ave., St. Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel: 314-362-1916, Fax: 314-747-2692."

ORIGIN

Query Match 88.2%; Score 529.8; DB 7; Length 572;
Best Local Similarity 99.3%; Pred. No. 6.8e-136;
Matches 553; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

```

QY 1 CTTGACGCGCCGGGCGCGTGC-GAAGCTCCCGAGTGGCCACAGATGCTTTAATTC 59
DB 17 CTTGACGCGCGCGGCGCGTGCAGGCACTCTCCAGTGGCCACAGATGCTTTAATTC 76
QY 60 AAAACCTGGCGCTTACAGGACAAAAGAGATCTTGGTAAATGCTGTTCAATCATCG 119
DB 77 AAAACCTGGCGCTTACAGGACAAAAGAGATCTTGGTAAATGCTGTTCAATCATCG 136
QY 120 CCACCACTTAATAGACGACAGATGAGAGTGTGGATAGACTCTACAGAGTACCA 179
DB 137 CCACCACTTAATAGACGACAGATGAGAGTGTGGATAGACTCTACAGAGTACCA 196
QY 180 GGGAGTACACCAAAACAAAGAGAGGCGAGAGATCTCAAGAACCTCATCAAGACG 239
DB 197 GGGAGTACACCAAAACAAAGAGAGGCGAGAGATCTCAAGAACCTCATCAAGACG 256
QY 240 TCATCAGCTGCGCACTTTTATAGGAATTAATCACTTAATCAAGATGAGTATGGA 299

```

Db 257 TCATCAAGCTGGCCATCTTTATAGAAATATCATGTTTATCAAGATGACGTACATGA 316
QY 300 TGGAGAAATTTAAGAGAAAGTTTCATCATGCTTGTCTATGACCGGTGTGATTCATCAGG 359
Db 317 TGGAGAAATTTAAGAGAAAGTTTCATCATGCTTGTCTATGACCGGTGTGATTCATCAGG 376
QY 360 TGGATTATACCTTTGACCGGAAATGTGTATTCACGAGCGCTTTAAATGAATGAGAGAGATGC 419
Db 377 TGGATTATACCTTTGACCGGAAATGTGTATTCACGAGCGCTTTAAATGAATGAGAGAGATGC 436
QY 420 TGCACCAATATCATTCAGCGCCACCTCATCTGCCAAGTCACATGAGCGGTTAAATATGCT 479
Db 437 TGCACCAATATCATTCAGCGCCACCTCATCTGCCAAGTCACATGAGCGGTTAAATATGCT 496
QY 480 TTGATCATTTTTCAGATTTGAAATTTTGGCTGCTGTGTAATTCCTTTTGGAAATTTTA 539
Db 497 TTGATCATTTTTCAGATTTGAAATTTTGGCTGCTGTGTAATTCCTTTTGGAAATTTTA 555
QY 540 AACCCCACTTACAAAA 556
Db 556 AACCCCACTTACAAAA 572

RESULT 35
AV713586 697 bp mRNA linear EST 11-OCT-2000
LOCUS AV713586 DCB Homo sapiens cDNA clone DCBBB11 5', mRNA sequence.
DEFINITION AV713586
ACCESSION AV713586
VERSION AV713586.1 GI:10795103
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 697)
XU, X., GU, J., LIU, F., QU, J., ZHAO, M., LI, Y., HUANG, Q., ZHOU, J.,
SONG, H., GU, Y., YANG, Y., GAO, G., XIAO, H., LI, N., QIAN, B., GAO, X.,
CHENG, Z., XU, S., GU, W., TU, Y., DIA, J., FU, G., REN, S., ZHONG, M.,
LU, G., CHENG, Z. and HAN, Z.
Homo sapiens cDNA DCB clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn
This clone is available at CHGC in Shanghai.

FEATURES
source
1..697
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DCBBB11"
/cell_type="dendritic cells"
/dev_stage="mature"
/lab_host="BM25.8"
/clone_lib="DCB"
/note="Vector: pTriplEx2; Site_1: sf11a; Site_2: sf11b"

ORIGIN
Query Match 88.1%; Score 529.4; DB 1; Length 697;
Best Local Similarity 98.9%; Pred. No. 9.3e-136;
Matches 533; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 36 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTACGACACAAAGAAATCTTG 95
Db 159 TGGCCACAGATGCTTTAATTCAGAAACCTGGCCGTTACGACACAAAGAAATCTTG 218
QY 96 GTAAATGCTGTCATATCGCCACGACCTTAATATAGACACAGAGATGAGAGTGC 155
Db 219 GTAAATGCTGTCATATCGCCACGACCTTAATATAGACACAGAGATGAGAGTGC 278

QY 156 TGGATGAGCTTACAGAGTGACGAGGAGTACACCCAAAAAGAGAGGCGAGAGAGA 215
Db 279 TGGATGAGCTTACAGAGTGACGAGGAGTACACCCAAAAAGAGAGGCGAGAGAGA 338
QY 216 TCATCAAGAACCTCATCAAGACAGTCATCAAGGTGGCCATTCTTTATAGGAATTAACAGT 275
Db 339 TCATCAAGAACCTCATCAAGACAGTCATCAAGGTGGCCATTCTTTATAGGAATTAACAGT 398
QY 276 TTAATCAAGATGAGCTAGCATTTGATGAGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGA 335
Db 399 TTAATCAAGATGAGCTAGCATTTGATGAGAGAAATTTAAGAAAGAAAGTTTCATCAGCTTGA 458
QY 336 TGAACGCTGTCAGTTTCATCAGTGAGATTAATTCCTTTGACCGGAATGTTATTCAGAGC 395
Db 459 TGAACGCTGTCAGTTTCATCAGTGAGATTAATTCCTTTGACCGGAATGTTATTCAGAGC 518
QY 396 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGGCGCCACTCAGTCCAGAT 455
Db 519 TGTAAATGAATGACAGAGATGCTGACCAAAATCATTCAGGCGCCACTCAGTCCAGAT 578
QY 456 CACATGAGCGGTTAATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 515
Db 579 CACATGAGCGGTTAATATGCTTTGATCATTTTTCAGATTGTGAATTTTGGCTGCT 638
QY 516 TGTAAATGCTTTTGGAAATTTTAAACCCCACTTACAAAAACTATGTGATGATCAAC 574
Db 639 TGTAAATGCTTTTGGAAATTTTAAACCCCACTTACAAAAACTATGTGATGATCAAC 697

RESULT 36
BG434534 739 bp mRNA linear EST 14-MAR-2001
LOCUS BG434534
DEFINITION 602506586F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:460317 5',
mRNA sequence.
ACCESSION BG434534
VERSION BG434534.1 GI:13341040
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 739)
NIH-MGC http://mhc.nci.nih.gov/.
TITLE NIH-MGC
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: CLOUTech Laboratories, Inc.
CDNA Library Preparation: CLOUTech Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LNCM1338 row: k column: 22
High quality sequence stop: 717.
Location/Qualifiers
1..739
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:460317"
/lab_host="PHIOB (T1 phage-resistant)"
/clone_lib="NIH_MGC_79"
/note="Organ: Placenta; Vector: pDMR-LIB (Clontech);
Site_1: Sf11 (ggccggccggcc); Site_2: Sf11
(ggccatcatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGCCCATTTATGGCC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGACGACATG-dT(30)BH-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.3
kb (range 0.5-4.0 kb). 15/15 colonies contained inserts

ORIGIN

by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

Query Match 87.8%; Score 527.6; DB 4; Length 739;

Best Local Similarity 97.7%; Pred. No. 3e-135;

Matches 588; Conservative 0; Mismatches 9; Indels 5; Gaps 5;

```
1 CTTGACGCTCCGCGCCGCTGCG-CGACTCTCTCCGATGCGCCACAGATGCTTTAATCCA 59
|||
76 CTTGACGCTCCGCGCCGCTGCGCCGCTCTCCGATGCGCCACAGATGCTTTAATCCA 135
|||
60 AAAACCTGCGCTTTCAGGACAAAAGATCTTGGTAAATGGTCCAAATTCATCG 119
|||
136 AAAACCTGCG-CGTTTCAGGACAAAAGATCTTGGTAAATGGTCCAAATTCATCG 194
|||
120 CCACCACTTAAATAGACACACAGATGAGGTGCTGATGAGCTCTACAGAGTACCA 179
|||
195 CCACCACTTAAATAGACACACAGATGAGGTGCTGATGAGCTCTACAGAGTACCA 254
|||
180 GGAAGTACACCCAAAACAAGAGAGGACAGAAAGATCTCAAGAACTCATCAACAG 239
|||
255 -GAGTACACCCAAAACAAGAGAGGACAGAAAGATCTCAAGAACTCATCAACAG 313
|||
240 TCATCAAGCTGCGCATTTCTTATAGGAATATCAGTTTAAATCAAGATGAGCTGATGA 299
|||
314 TCATCAAGCTGCGCATTTCTTATAGGAATATCAGTTTAAATCAAGATGAGCTGATGA 373
|||
300 TGAAGAAATTTAAGAAAGATTCATCAGCTTGTATGACCGTGTACAGTTTCATCAG 359
|||
374 TGAAGAAATTTAAGAAAGATTCATCAGCTTGTATGACCGTGTACAGTTTCATCAG 433
|||
360 TGAATATACCTTTCAGCCGAAATGCTTATCCAGGCTGTTAAATGATCAAGAGATGC 419
|||
434 TGAATATACCTTTCAGCCGAAATGCTTATCCAGGCTGTTAAATGATCAAGAGATGC 493
|||
420 TGAACCAATTCATTCAGCCGCACTCAGTGCAGATCAGATGAGCGGTTAATATGCT 479
|||
494 TGAACCAATTCATTCAGCCGCACTCAGTGCAGATCAGATGAGCGGTTAATATGCT 553
|||
480 TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGTATATATCTTTTGGGAATTTTA 539
|||
554 TTGATCATTTTTCAGATGTGAATTTTGGCTGCTGATATATCTTTTGGGAATTTTA 612
|||
540 AACCCCACTTACAAAACATATGATGATGATCAACAATGTTGATGAGAGAACTAT 599
|||
613 AACCCCACTTACAAAACATATGATGATGATCAACAATGTTGATGAGAGAACTAT 671
|||
600 GA 601
|||
672 GA 673
```

RESULT 37

CR71358 692 bp mRNA linear EST 23-SEP-2004

LOCUS DKEP46910535_r1_469 (synonym: pkid1) Pongo pygmaeus cDNA clone

DEFINITION DKF2P46910535_5', mRNA sequence.

ACCESSION CR71358

VERSION CR71358.1 GI:52614631

KEYWORDS EST.

SOURCE Pongo pygmaeus (orangutan)

ORGANISM Pongo pygmaeus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.

REFERENCE 1 (bases 1 to 692)

Pousetka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,

Mewes, H.W., Well, B., Amid, C., Osanger, A., Fobo, G., Han, M. and

Wiemann, S.

Pongo pygmaeus mRNA (Pousetka, A., Albert, R., Moosmayer, P., et al.)

Unpublished (2004)

COMMENT CONTACT: MIPS

FEATURES

source

1..692
location/Qualifiers

/organism="Pongo pygmaeus"

/mol_type="mRNA"

/db_xref="taxon:9600"

/clone="DKF2P46910535"

/tissue_type="kidney"

/dev_stage="adult"

/lab_host="DH10B"

/clone_1b="469 (synonym: pkid1)"

/note="Vector: pSPORT1_sfi; Site_1: SfiIa; Site_2: SfiIb"

ORIGIN

Query Match 87.6%; Score 526.6; DB 7; Length 692;

Best Local Similarity 99.1%; Pred. No. 5.5e-135;

Matches 540; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

```
36 TGGCCACAGATGCTTTAATTCAAAACCTGGCGGTCAGGACAAAAGAGATCTTGG 95
|||
149 TGGCCACAGATGCTTTAATTCAAAACCTGGCGGTCAGGACAAAAGAGATCTTGG 208
|||
96 GTAAATGCTGTCCAATTCATGCGCACCCCTTAATAGACACACAAATAGTGAAGTGC 155
|||
209 GTAAATGCTGTCCAATTCATGCGCACCCCTTAATAGACACACAAATAGTGAAGTGC 268
|||
156 TGAATGAGCTTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTAC 215
|||
269 TGAATGAGCTTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTAC 328
|||
216 TCATCAAGAACCTCATCAAGAACCTCATCAAGAACCTCATCAAGAACCTCATCAAG 275
|||
329 TCATCAAGAACCTCATCAAGAACCTCATCAAGAACCTCATCAAGAACCTCATCAAG 388
|||
276 TTAATCAAGATGAGTAGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 335
|||
389 TTAATCAAGATGAGTAGATGATGAGAAATTTAAGAAAGATTCATCAGCTTGCTA 448
|||
336 TGAACGCTGCTCAGTTTCATCAGTGTGATTAATCCTTGAACCGGAATGTTATCCAG 395
|||
449 TGAACGCTGCTCAGTTTCATCAGTGTGATTAATCCTTGAACCGGAATGTTATCCAG 508
|||
396 TGTAAATGAATGACAGAGATGCTGACCAATCATTCAGGCGCACTCATGCAAGT 455
|||
509 TGTAAATGAATGACAGAGATGCTGACCAATCATTCAGGCGCACTCATGCAAGT 568
|||
456 CACATGACGCGGTATATATGCTTTCATCAGATTTTCAGATTTGATTTTGGCTGCT 515
|||
569 CACATGACGCGGTATATATGCTTTCATCAGATTTTCAGATTTGATTTTGGCTGCT 628
|||
516 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATATGATGATGATCA 575
|||
629 TGTATATCTTTTGGGAATTTTAAACCCCACTTACAAAACATATGATGATGATCA 687
|||
576 AATG 580
|||
688 AATG 692
```

RESULT 38

BG534704

LOCUS

DEFINITION

ACCESSION

BG534704

685 bp mRNA linear EST 03-APR-2001

602553867F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:463672 5',

mRNA sequence.

BG534704

VERSION BG534704.1 GI:13525246
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 685)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Clontech Laboratories, Inc.
 cDNA Library Preparation: Clontech Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
<http://image.llnl.gov>
 Plate: L10M1466 row: e column: 17
 High quality sequence clone: 683.
 Location/Qualifiers
 1..685
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:463672"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_1lb="NIH MGC 77"
 /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1:
 SfiI (ggcgccctcgccc); Site 2: SfiI (ggcgccatagccc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGCCCATTTGCGC-3' and 3' adaptor sequence
 5'-ATTCTAGAGCCGAGCGCGGCAGCATG-dT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.9
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH-MGC library."

QY	475	TGCTTTGATGATCTTTTCAGATTGTGAAATTTTGGCGCGCCCTTGATATGATCTTTTGGGAA	534
Db	420	TGTGTTGATATATTTTCAGATTGTGGA-TTTTGGCGCCCTTGATATATCTTTTGGGAA	478
QY	535	TTTTAAACCCCACTTACAAAACACTATGTGATGTGATACAAAATGTGTGATGAAGAA	594
Db	479	TTTTAAACCCCACTTACAAAACACTATGTGATGTGATACAAAATGTGTGATGAAGAA	538
QY	595	CATATGA	601
Db	539	CATATGA	545
RESULT 39			
LOCUS	CV027255		
DEFINITION	CV027255	5454 Full Length cDNA from the Mammalian Gene Collection Homo sapiens cDNA 5' similar to BC005352, mRNA sequence.	
ACCESSION	CV027255		
VERSION	CV027255.1	GI:51485291	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.		
REFERENCE	1 (bases 1 to 593)		
AUTHORS	Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S., Driscot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O., Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T., Simmons,B., Sequerra,R., Bosak,S., Doucette-Stamm,L., Le Peuch,C., Vandenhaute,J., Guisic,M.E., Alcala,J.S., Hill,D.E. and Vidal,M.		
TITLE	Human ORFome Version 1.1: a Platform for Reverse Proteomics		
JOURNAL	Genome Res. (2004) In press		
COMMENT	Contact: Vidal M		
	Marc Vidal Laboratory		
	Dana Farber Cancer Institute		
	1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA		
	Tel: 617 632 5180		
	Fax: 617 632 5739		
	Email: Marc.Vidal@dfci.harvard.edu		
	ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF results from a PCR reaction using an MGC full-length cDNA as template DNA and ORF specific primers		
	PCR primers		
	FORWARD: ATGCATCTCCAGACGAAGAATCCA		
	BACKWARD: CATATGTCTTTCATCAACAATTT		
	Insert Length: 593 Std Error: 78.00		
	Plate: 11028 row: 11 column: H		
	Seq primer: ACTGCGCGTGTTCATCAAGCTGTGACTGGAAAC		
	High quality sequence start: 101		
	High quality sequence stop: 592		
	POLYA=No.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..593		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="mixed"		
	/clone_lib="Full Length cDNA from the Mammalian Gene Collection"		
	/note="Vector: mixed; The ORFs were PCR amplified from the MGC (Mammalian Gene Collection) as of April 2004 and cloned by recombinational Gateway cloning into pDONR223 Donor Vector. Reference : MGC (Mammalian Gene Collection) Program Team, Generation and Initial Analysis of more than 15,000 Full-length Human and Mouse cDNA Sequences. PNAS, 2002, 99(26), 16899-16903"		
ORIGIN			
Query Match	86.4%	Score 519.2	DB 7
Best Local Similarity	99.0%	Pred.No. 6.1e-133	Length 593;
Matches 547;	Conservative 0;	Mismatches 8;	Indels 3;
			Gaps 2;


```
QY 36 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTCAGGCACAAAAGAGATCTGG 95
DB 32 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTCAGGCACAAAAGAGATCTGG 91
QY 96 GTAAAAATGCTGTCACAAATCATCGCCACACCTTAATAGACACACAGTAGTAGTGC 155
DB 92 GTAAAAATGCTGTCACAAATCATCGCCACACCTTAATAGACACACAGTAGTAGTGC 151
QY 156 TGGATGAGCTTACAGAGTACAGGAGGATGACACCCAAAACAAGAGAGGACAGAGA 215
DB 152 TGGAGAGCTTACAGAGTACAGGAGGATGACACCCAAAACAAGAGAGGACAGAGA 211
QY 216 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGAAATATCACT 275
DB 212 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGAAATATCACT 271
QY 276 TTAATCAAGATGAGCTAGCTAGTGAAGAAATTTAAGAAAGAAAGTTCACTCACTTGTCTA 335
DB 272 TTAATCAAGATGAGCTAGCTAGTGAAGAAATTTAAGAAAGAAAGTTCACTCACTTGTCTA 331
QY 336 TGAACGCTGCTCACTTCCATCAGTGGATTATACCTTTGACCGGAATGTTATCCAGGC 395
DB 332 TGAACGCTGCTCACTTCCATCAGTGGATTATACCTTTGACCGGAATGTTATCCAGGC 391
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCCACTCACTGCCAAGT 455
DB 392 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCCACTCACTGCCAAGT 451
QY 456 CACATGACAGGCTTAATTAATGCTTTGATCATTTTTCAGATGTGAAATTTTGGCTGCT 515
DB 452 CACATGACAGGCTTAATTAATGCTTTGATCATTTTTCAGATGTGAAATTTTGGCTGCT 511
QY 516 TGTATTAATCCTTTGGGAATTTTAAACCGCACTTCAAAAACATATGATGATGAATCA 575
DB 512 TGTATTAATCCTTTGGGAATTTTAA--CCCACTTACAAAACATATGATGATGAATCAAC- 568
QY 576 AAATGTTGGATGAAGAGA 593
DB 569 AAATGTTGGATGAAGAAA 586

RESULT 40
CD520319 757 bp mRNA linear EST 06-JUN-2003
LOCUS AGECONRT 14360077 NIH MGC 191 Homo sapiens cDNA clone
DEFINITION IMAGE:30410454 5', mRNA sequence.
ACCESSION CD520319
VERSION CD520319.1 GI:31452037
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 757)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: dsgerha@mail.nih.gov
Tissue Procurement: Narayan Bhat
cDNA Library Preparation: CLOUTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: NDCM200 row: k column: 07
High quality sequence stop: 563.
Location/Qualifiers
1..757
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30410454"
/libsize_type="Pooled"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NIH MGC 191"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattcggcc);
Site 2: SfiI (ggccgctggcc); Library is oligo-dT primed
and directionally cloned. PMBC - Peripheral Blood
Mononuclear Cells. RNA was pooled from 3/6hour stimulation
with PMA adn Ionomycin. 5' and 3' adaptors were used in
cloning as follows: 5' adaptor sequence:
5'-CACGGCATTATGCGC-3' and 3' adaptor sequence:
5'-ATTCTAAGAGCGGAGCGGCGACAGTGTG-3' (30-BN-3'
C, or G and N = A, C, G, or T). Average insert size 1.69
kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC library."

ORIGIN
Query Match 85.5%; Score 513.6; DB 6; Length 757;
Best Local Similarity 98.7%; Pred. No. 2.3e-131;
Matches 539; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 36 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTCAGGCACAAAAGAGATCTGG 95
DB 32 TGGCCACAGATGCTTTAATTCACAAAACCTGGCCGCTTCAGGCACAAAAGAGATCTGG 91
QY 96 GTAAAAATGCTGTCACAAATCATCGCCACACCTTAATAGACACACAGTAGTAGTGC 155
DB 92 GTAAAAATGCTGTCACAAATCATCGCCACACCTTAATAGACACACAGTAGTAGTGC 151
QY 156 TGGATGAGCTTACAGAGTACAGGAGGATGACACCCAAAACAAGAGAGGACAGAGA 215
DB 152 TGGAGAGCTTACAGAGTACAGGAGGATGACACCCAAAACAAGAGAGGACAGAGA 211
QY 216 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGAAATATCACT 275
DB 212 TCATCAAGAACCTCATCAAGACAGTCATCAAGCTGGCCATTCTTTATAGAAATATCACT 271
QY 276 TTAATCAAGATGAGCTAGCTAGTGAAGAAATTTAAGAAAGAAAGTTCACTCACTTGTCTA 335
DB 272 TTAATCAAGATGAGCTAGCTAGTGAAGAAATTTAAGAAAGAAAGTTCACTCACTTGTCTA 331
QY 336 TGAACGCTGCTCACTTCCATCAGTGGATTATACCTTTGACCGGAATGTTATCCAGGC 395
DB 332 TGAACGCTGCTCACTTCCATCAGTGGATTATACCTTTGACCGGAATGTTATCCAGGC 391
QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCCACTCACTGCCAAGT 455
DB 392 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCCACTCACTGCCAAGT 451
QY 456 CACATGACAGGCTTAATTAATGCTTTGATCATTTTTCAGATGTGAAATTTTGGCTGCT 515
DB 452 CACATGACAGGCTTAATTAATGCTTTGATCATTTTTCAGATGTGAAATTTTGGCTGCT 511
QY 516 TGTATTAATCCTTTGGGAATTTTAAACCGCACTTCAAAAACATATGATGATGAATCA 575
DB 512 TGTATTAATCCTTTGGGAATTTTAA--CCCACTTACAAAACATATGATGATGAATCAAC- 568
QY 576 AAATGTTGGATGAAGAGA 593
DB 569 AAATGTTGGATGAAGAAA 586

QY 396 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCCACTCACTGCCAAGT 455
DB 467 TGTAAATGAATGACAGAGATGCTGCACCAATCATTCAGGCGCCACTCACTGCCAAGT 456
QY 456 CACATGACAGGCTTAATTAATGCTTTGATCATTTTTCAGATGTGAAATTTTGGCTGCT 515
DB 527 CACATGACAGGCTTAATTAATGCTTTGATCATTTTTCAGATGTGAAATTTTGGCTGCT 586
QY 516 TGTATTAATCCTTTGGG--AATTTAAACCCCACTTCAAAAACATATGATGA-TGATATCA 572
DB 587 TGTATTAATCCTTTGGGAAATTTTAAACCCCACTTCAAAAACATATGATGATGATATCA 646
QY 573 ACAA 578
DB 647 ACAA 652

RESULT 41
CD723203 588 bp mRNA linear EST 26-JUN-2003
LOCUS oj19b07.y1 Human lacrimal gland, unamplified: oj Homo sapiens cDNA
DEFINITION clone oj19b07 5', mRNA sequence.
ACCESSION CD723203
```

VERSION CD223203.1 GI:32274051
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 588)
 AUTHORS Dickinson, D., Laurie, G. and Wistow, G.
 TITLE Expressed sequence tag analysis of human lacrimal gland
 JOURNAL Unpublished (2002)
 COMMENT Contact: Wistow, G.
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: graeme@helix.nih.gov
 Plate: 19 row: b column: 07
 Seq primer: M13RPI reverse primer (ABI).
 Location/Qualifiers
 1..588
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="oj19p07"
 /tissue_type="lacrimal gland"
 /dev_stage="Adult"
 /lab_host="EMDH10B"
 /note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted from 2 human lacrimal glands. A directionally cloned cDNA library in the pCMVSPORT6 vector (Life Technologies) was constructed at Bioserve Biotechnology (Laurie MD) essentially following the protocols of which are contained in the manufacturer's instruction manual (http://www.lifetech.com/). First strand synthesis was carried out using a Not I primer-adaptor (5'-GACTAGTCTAGATCGCGAGCGGCCGCTT)15-3'. EST analysis was performed on the unamplified library at the NIH Intramural Sequencing Center (NISC)."
 ORIGIN

Query Match 84.8%; Score 509.4; DB 6; Length 588;
 Best Local Similarity 99.8%; Pred. No. 3.2e-130;
 Matches 510; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 91 CTTGGGTAAGATGATGCAATTCATCCGACACCTTATAGAGACCAAGTAGTGA 150
 DB 1 CTTGGGTAAGATGATGCAATTCATCCGACACCTTATAGAGACCAAGTAGTGA 60
 QY 151 GGTGCTGATGAGCTCTACAGATGACCAAGGATACACCAAGCAAGAGGACAGA 210
 DB 61 GGTGCTGATGAGCTCTACAGATGACCAAGGATACACCAAGCAAGAGGACAGA 120
 QY 211 GAAGATCATCAAGACCTCTCAAGACATCATCAAGCTGCGCATCTTTATAGATAA 270
 DB 121 GAAGATCATCAAGACCTCTCAAGACATCATCAAGCTGCGCATCTTTATAGATAA 180
 QY 271 TCAGTTTATCAAGATGAGTACATTTGATGAGAAATTAAGAAAGTTATAGACT 330
 DB 181 TCAGTTTATCAAGATGAGTACATTTGATGAGAAATTAAGAAAGTTATAGACT 240
 QY 331 TCGTATGACCGTGTGATGCTTCATCAGTGTGATTAACCTTTGACCGAGATGTTATC 390
 DB 241 TCGTATGACCGTGTGATGCTTCATCAGTGTGATTAACCTTTGACCGAGATGTTATC 300
 QY 391 CAGGCTTTAAAGATGACAGAGATGCTGCACCAATCATTCAGGCGCACTCACTGC 450
 DB 301 CAGGCTTTAAAGATGACAGAGATGCTGCACCAATCATTCAGGCGCACTCACTGC 360
 QY 451 CAAGTCAATGACGCGGTTATATGCTTTGATCAATTTTTCAGATGTAATTTTGGC 510

DB 361 CAAGTCAATGACGCGGTTATATGCTTTGATCAATTTTTCAGATGTAATTTTGGC 420
 QY 511 TCGCTTGTATATCTCTTTTGGGATTTTAAACCCCACTTACAAAATCTATGATGTTAT 570
 DB 421 TCGCTTGTATATCTCTTTTGGGATTTTAAACCCCACTTACAAAATCTATGATGTTAT 480
 QY 571 CAACAAATGTTGATGAGAGAGAAATATGA 601
 DB 481 CAACAAATGTTGATGAGAGAGAAATATGA 511

RESULT 42
 BP345938
 LOCUS BP345938
 DEFINITION BP345938 Sugano cDNA library, synovial membrane Homo sapiens CDNA
 clone SYN08154, mRNA sequence.
 ACCESSION BP345938
 VERSION BP345938.1 GI:52275919
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 583)
 AUTHORS Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
 Mizushima-Sugano, J., Nakai, K. and Sugano, S.
 TITLE Sequence comparison of human and mouse genes reveals a homologous
 block structure in the promoter regions
 JOURNAL Genome Res. 14 (9), 1711-1718 (2004)
 COMMENT Contact: Yutaka Suzuki
 Department of Virology
 Institute of Medical Science, University of Tokyo
 4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
 Email: yusuzuki@ims.u-tokyo.ac.jp.
 Location/Qualifiers
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="SYN08154"
 /tissue_type="synovial membrane"
 /clone_lib="Sugano cDNA library, synovial membrane"
 ORIGIN

Query Match 82.8%; Score 497.8; DB 5; Length 583;
 Best Local Similarity 99.4%; Pred. No. 5.3e-127;
 Matches 510; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 1 CTTACGCTCCCGGCGCGCGTCCG-CGACTCTCTCCATGCGCACAGATGCTTTAATTCGA 59
 DB 71 CTTACGCTCCCGGCGCGCGTCCGCGCACCTCCGATGGCCACAGATGCTTTAATTCGA 130
 QY 60 AAAACCTGGCCGTTGAGGACCAAGAGATCTGGGTAAATGGGTCCAATTCATCG 119
 DB 131 AAAACCTGGCCGTTGAGGACCAAGAGATCTGGGTAAATGGGTCCAATTCATCG 190
 QY 131 AAAACCTGGCCGTTGAGGACCAAGAGATCTGGGTAAATGGGTCCAATTCATCG 190
 DB 120 CCACCACTTATATAGACGACACAGTAGTAGGTGCTGATGAGCTCTACAGAGTACCA 179
 QY 120 CCACCACTTATATAGACGACACAGTAGTAGGTGCTGATGAGCTCTACAGAGTACCA 179
 DB 191 CCACCACTTATATAGACGACACAGTAGTAGGTGCTGATGAGCTCTACAGAGTACCA 250
 QY 180 GGGAGTACACCCAAAGCAAGAGAGGCGAGAGATCATCAAGAACTTCATTAAGACG 239
 DB 251 GGGAGTACACCCAAAGCAAGAGAGGCGAGAGATCATCAAGAACTTCATTAAGACG 310
 QY 240 TCATCAAGCTGGCCATCTTTATAGAAATATCAAGTTATCAAGATGAGCTAGCTTGA 299
 DB 311 TCATCAAGCTGGCCATCTTTATAGAAATATCAAGTTATCAAGATGAGCTAGCTTGA 370
 QY 300 TGGAGAAATTTAAGAAAGATTCATCAGCTTGTATGACCGTGTGAGTTTCATGACG 359
 DB 371 TGGAGAAATTTAAGAAAGATTCATCAGCTTGTATGACCGTGTGAGTTTCATGACG 430
 QY 360 TGGATTAATCTTTGACCGGAATGTGTTATCCAGGCTGTAAATGAATGACAGAGATGC 419

Db 431 TGGATATACCTTTGACCGGAATGTGTATCCAGGCTGTTAATGATGAGAGATGC 490
Qy 420 TGCACCAATTCATTCAGCGCCACCTCACGCGCAAGTCACATGACGGGTAAATGCT 479
Db 491 TGCACCAATTCATTCAGCGCCACCTCACGCGCAAGTCACATGACGGGTAAATGCT 550
Qy 480 TGTATCATTTTTCAGATGTGGAATTTTGGCTG 512
Db 551 TGTATCATTTTTCAGATGTGGAATTTTGGCTG 583

RESULT 43
LOCUS CR771295 605 bp mRNA linear EST 23-SEP-2004
DEFINITION DKFZp469M0735_r1_469 (synonym: pkid1) Pongo pygmaeus CDNA clone
VERSION DKFZp469M0735_5', mRNA sequence.
KEYWORDS CR771295
SOURCE CR771295.1 GI:52614568
ORGANISM Pongo pygmaeus (orangutan)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pongo.
REFERENCE Pousetka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Wiemann, S., Mewes, H. W., Weill, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.
Pongo pygmaeus mRNA (Pousetka, A., Albert, R., Moosmayer, P., et al.)
JOURNAL Unpublished (2004)
COMMENT Contact: MIPS
MIPS
Ingolsteader Landstr. 1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert. Clone from S. Wiemann,
Molecular Genome Analysis, German Cancer Research Center (DKFZ);
Email: s.wiemann@dkfz-heidelberg.de; any. Please contact RZPD for
ordering:
http://www.rzpd.de/cgi-bin/product/ci.cgi?cloneID=DKFZp469M0735
Further information about the clone and the sequencing project is
available at http://mips.gsf.de/projects/cdna/.

FEATURES
source
1..605
/organism="Pongo pygmaeus"
/mol_type="mRNA"
/db_xref="taxon:9600"
/clone="DKFZp469M0735"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="469 (synonym: pkid1)"
/note="Vector: pSPORT1_sfi; site_1: sfi1A; site_2: sfi1B"

ORIGIN
Query Match 82.6%; Score 496.6; DB 7; Length 605;
Best Local Similarity 99.2%; Pred. No. 1.2e-126;
Matches 499; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 36 TGGCCACATATGCTTTAATTCCTGCGCTTCAGGACCAAAAGAAATTTGG 95
Db 103 TGGCCACATATGCTTTAATTCCTGCGCTTCAGGACCAAAAGAAATTTGG 162
Qy 96 GTAATAATGCTTCATTCATTCGACCACTTAATAGACACAAAGTAGAGAGTGC 155
Db 163 GTAATAATGCTTCATTCATTCGACCACTTAATAGACACAAAGTAGAGAGTGC 222
Qy 156 TGGATAGCTTACAGAGTACCGAGGAGTACCCCAAAACAAAGAGAGGACAGAGA 215
Db 223 TGGATAGCTTACAGAGTACCGAGGAGTACCCCAAAACAAAGAGAGGACAGAGA 282
Qy 216 TGTATCAAGAACTTCATCAAGAGTATCATGATGCGCATTTTATAGAAATATCACT 275
Db 283 TGTATCAAGAACTTCATCAAGAGTATCATGATGCGCATTTTATAGAAATATCACT 342

Qy 276 TTAATCAAGATAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCATCAGTGTCTA 335
Db 343 TTAATCAAGATAGCTAGCATTTGATGAGAAATTTAAGAGAAAGTTCATCAGTGTCTA 402
Qy 336 TGAACCGTGTCACTTTCCATCAGGTGATTTATACCTTGAACCGGATGTTATCCAGGC 395
Db 403 TGAACCGTGTCACTTTCCATCAGGTGATTTATACCTTGAACCGGATGTTATCCAGGC 462
Qy 396 TGTAAATGAATGACAGAGATGTCGACCAATCATTCAGCGCCACCTCACTGCCAAGT 455
Db 463 TGTAAATGAATGACAGAGATGTCGACCAATCATTCAGCGCCACCTCACTGCCAAGT 522
Qy 456 CACATGACCGGTTAATATGCTTTGATCATTTTTCAGATTTGAAATTTTGGTGCT 515
Db 523 CACATGACCGGTTAATATGCTTTGATCATTTTTCAGATTTGATTTTGGTGCT 582
Qy 516 TGTATATCTTTTGGCAATTTT 538
Db 583 TGTATATCTTTTGGCAATTTT 605

RESULT 44
LOCUS BT60235 744 bp mRNA linear EST 25-SEP-2001
DEFINITION 603045171P1 NIH_MGC_116 Homo sapiens CDNA clone IMAGE:5185720 5',
mRNA sequence.
ACCESSION BT60235
VERSION BT60235.1 GI:15751813
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgs.mci.nih.gov/.
AUTHORS NIH-MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: L1AM11463 row: m column: 17
High quality sequence stop: 743.
Location/Qualifiers
1..744
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5185720"
/lab_host="DH10B"
/clone_lib="NIH MGC 116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; site_1: Not; site_2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2
stomachs, 62 yo male and 40 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 82.3%; Score 494.6; DB 4; Length 744;
Best Local Similarity 97.5%; Pred. No. 4.4e-126;
Matches 555; Conservative 0; Mismatches 9; Indels 5; Gaps 5;

QY 36 TGCCACAGATGCTTTAATTCACAAAACCTGGCCGTCAGGACCAAAAGAGATCTGG 95
 |||||
 DB 175 TGCCACAGATGCTTTAATTCACAAAACCTGGCCGTCAGGACCAAAAGAGATCTGG 234
 |||||
 QY 96 GTAAATGCTGTCACCAATCCATGCGCCACCACTTAATAGACACACAAGTAGAGTGCG 155
 |||||
 DB 235 GTAAATGCTGTCACCAATCCATGCGCCACCACTTAATAGACACACAAGTAGAGTGCG 234
 |||||
 QY 156 TGGATGAGCTTACAGAGTACCAAGGAGTACCAACCAAGAGAGGAGGAGAGAGA 215
 |||||
 DB 295 TGGATGAGCTTACAGAGTACCAAGGAGTACCAACCAAGAGAGGAGGAGAGAGA 354
 |||||
 QY 216 TCATCAAGAACCTCAAGACAGTACATCAAGCTGGCCATCTTTATAGAAATATCACT 275
 |||||
 DB 355 TCATCAAGAACCTCAAGACAGTACATCAAGCTGGCCATCTTTATAGAAATATCACT 414
 |||||
 QY 276 TTAATCAAGATGAGTACAGTATGATGAGAAATTTAA-GAAGAAAGTTCAGCTTGC 334
 |||||
 DB 415 TTAATCAAGATGAGTACAGTATGATGAGAAATTTAA-GAAGAAAGTTCAGCTTGC 474
 |||||
 QY 335 ATGACCGTGGTCACTTCCATCAAGTGGATTATACCTTTGACCGGAATGTTATCCAGG 394
 |||||
 DB 475 ATGACCGTGGTCACTTCCATCAAGTGGATTATACCTTTGACCGGAATGTTATCCAGG 534
 |||||
 QY 395 CTGTAAATGAATGCA-GAAGATGCTGCAACCAATCATTCAGCGCCACCTCACTGCCAA 453
 |||||
 DB 535 CTGTAAATGAATGCA-GAAGATGCTGCAACCAATCATTCAGCGCCACCTCACTGCCAA 594
 |||||
 QY 454 GTCACATGACGGGTTAATATATGCTTTCATTTTCAGTTTTCAGTTTTCAGTTTTCAGTTG 513
 |||||
 DB 595 GTCACATGACGGGTTAATATATGCTTTCATTTTCAGTTTTCAGTTTTCAGTTTTCAGTTG 652
 |||||
 QY 514 CTGTAAATGCTTTTGGAAATTTTAAACCCCACTTACAAAAAC-TATGATGATGATCA 572
 |||||
 DB 653 CTGTAAATGCTTTTGGAAATTTTAAACCCCACTTACAAAAAC-TATGATGATGATCA 712
 |||||
 QY 573 ACMAAATGTTGATGAAGAGACATATGA 601
 |||||
 DB 713 ACMAAATGTTGATGAAGAGACATATGA 741
 |||||
 RESULT 45
 BP433798 773 bp mRNA linear EST 30-DEC-2003
 LOCUS BP433798 full-length enriched swine cDNA library, adult lung Sus
 DEFINITION BP433798
 ACCESSION BP433798
 VERSION BP433798.1 GI:40423865
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 1 (bases 1 to 773)
 Uenishi,H., Eguchi,T., Suzuki,K., Sawazaki,T., Toki,D., Shinkai,H.,
 Okumura,N., Hamasima,N. and Awata,T.
 PEDF (Pig EST Data Explorer): construction of a database for ESTs
 derived from porcine full-length cDNA libraries
 Nucleic Acids Res. 32 (1), D484-D488 (2004)
 Contact: Hirohide Uenishi
 Animal Genome Laboratory, Genome Research Department
 National Institute of Agrobiological Sciences
 2 Ikenodai, Tsukuba, Ibaraki 305-8602, Japan
 Tel: +81-29-838-8627
 Fax: +81-29-838-8627
 Email: huenishi@affrc.go.jp
 EST project with full-length enriched cDNA libraries carried out in
 Animal Genome Research Program (Japan) by National Institute of
 Agrobiological Sciences and STAF-Institute
 Single pass sequencing of clones derived from oligo-capped cDNA
 library
 Vector sequences were eliminated by RepeatMasker version 2002/07/13
 and crossmatch version 0.990319

FEATURES
 Source
 1..773
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /clone="LNG010079A10"
 /isue_type="adult"
 /dev_stage="adult"
 /clone_1ib="full-length enriched swine cDNA library, adult
 lung"
 ORIGIN
 Query Match 81.4%; Score 489.2; DB 5; Length 773;
 Best Local Similarity 91.5%; Pred. No. 1.4e-124;
 Matches 518; Conservative 0; Mismatches 48; Indels 0; Gaps 0;
 QY 36 TGCCACAGATGCTTTAATTCACAAAACCTGGCCGTCAGGACCAAAAGAGATCTGG 95
 |||||
 DB 155 TGCCACAGATGCTTTAATTCACAAAACCTGGCCGTCAGGACCAAAAGAGATCTGG 214
 |||||
 QY 96 GTAAATGCTGTCACCAATCCATGCGCCACCACTTAATAGACACACAAGTAGAGTGCG 155
 |||||
 DB 215 GTAAATGCTGTCACCAATCCATGCGCCACCACTTAATAGACACACAAGTAGAGTGCG 274
 |||||
 QY 156 TGGATGAGCTTACAGAGTACCAAGGAGTACCAACCAAGAGAGGAGGAGAGAGA 215
 |||||
 DB 275 TGGATGAGCTTACAGAGTACCAAGGAGTACCAACCAAGAGAGGAGGAGAGAGA 334
 |||||
 QY 216 TCATCAAGAACCTCAAGACAGTACATCAAGCTGGCCATCTTTATAGAAATATCACT 275
 |||||
 DB 335 TCATCAAGAACCTCAAGACAGTACATCAAGCTGGCCATCTTTATAGAAATATCACT 394
 |||||
 QY 276 TTAATCAAGATGAGTACAGTATGATGAGAAATTTAA-GAAGAAAGTTCAGCTTGC 335
 |||||
 DB 395 TTAATCAAGATGAGTACAGTATGATGAGAAATTTAA-GAAGAAAGTTCAGCTTGC 454
 |||||
 QY 336 ATGACCGTGGTCACTTCCATCAAGTGGATTATACCTTTGACCGGAATGTTATCCAGG 395
 |||||
 DB 455 ATGACCGTGGTCACTTCCATCAAGTGGATTATACCTTTGACCGGAATGTTATCCAGG 514
 |||||
 QY 396 CTGTAAATGAATGCA-GAAGATGCTGCAACCAATCATTCAGCGCCACCTCACTGCCAA 455
 |||||
 DB 515 CTGTAAATGAATGCA-GAAGATGCTGCAACCAATCATTCAGCGCCACCTCACTGCCAA 574
 |||||
 QY 456 CACATGACGGGTTAATATGCTTTCATTTTCAGTTTTCAGTTTTCAGTTTTCAGTTTTCAGTTG 515
 |||||
 DB 575 CACATGACGGGTTAATATGCTTTCATTTTCAGTTTTCAGTTTTCAGTTTTCAGTTTTCAGTTG 634
 |||||
 QY 516 TGTATATCCTTTTGGAAATTTTAAACCCCACTTACAAAACTATGATGATGATCACT 575
 |||||
 DB 635 TGTATATCCTTTTGGAAATTTTAAACCCCACTTACAAAACTATGATGATGATCACT 694
 |||||
 QY 576 AAATGTTGATGAAGAGACATATGA 601
 |||||
 DB 695 AAATGTTGATGAAGAGACATATGA 720
 |||||

Search completed: Jul 28, 2005, 19:01:12
 Job time : 2905 secs

THIS PAGE BLANK (USPTO)